Chapter 2
Basic number theory

In this chapter we cover basic number theory and set up notations that will be used freely throughout the rest of the book. The chapter starts with the basic notions of divisibility and prime numbers with the goal of proving the Fundamental Theorem of Arithmetic, Theorem 2.19. We then prove the Chinese Remainder Theorem (Theorem 2.24), Fermat's Little Theorem (Theorem 2.26), Euler's Theorem (Theorem 2.31), discuss the basic properties of the totient function ϕ, and study polynomials modulo primes, digit expansions, and finally primitive roots. In the Notes at the end of the chapter, we talk about Euclid and his masterpiece the *Elements*; briefly discuss natural numbers and induction; review two standard cryptographic methods based on number theory; and finally, state Artin's conjecture for primitive roots.

2.1 Natural numbers, mathematical induction, and the Well-ordering Principle

The numbers $1, 2, 3, \ldots$ are called *natural numbers*, and we denote the set of all natural numbers by $\mathbb{N}$. A defining property of the set of natural number is the following:

Property 2.1 (Mathematical induction). Let $A \subset \mathbb{N}$ be such that

- $1 \in A$;
- $x \in A$ implies $x + 1 \in A$.

Then $A = \mathbb{N}$.

The set of natural numbers has the following fundamental property as well:

Property 2.2 (Well-ordering Principle). Every non-empty subset of the set of natural numbers has a smallest element.

For example, if we consider the subset of the set of natural numbers consisting of all even numbers, then the smallest element of this set is the number 2; or, if the

R. Takloo-Bighash, *A Pythagorean Introduction to Number Theory*,
Undergraduate Texts in Mathematics, https://doi.org/10.1007/978-3-030-02604-2_2

subset is the set of all multiples of 75, then the smallest element is 75. Intuitively, the Well-ordering Principle is true because the set of natural numbers does not go *all the way down*, though this is of course not a proof. In fact, the Well-ordering Principle is equivalent to mathematical induction.

Theorem 2.3. *The Well-ordering Principle is logically equivalent to mathematical induction.*

Proof. First we show that mathematical induction implies the Well-ordering Principle. Let P_n be the following statement: Every subset of $\mathbb{N}$ which contains a number x such that $x \leq n$ has a smallest element. Clearly P_1 is true, as in this case the subset will contain 1, and 1 will be the smallest element. So now suppose we know P_k is true, and we wish to show P_{k+1} is true. Suppose $A \subset \mathbb{N}$ is such that A contains some element x with $x \leq k+1$. If A contains some element y with $y \leq k$, then the validity of P_k implies that A must have a smallest element. So assume there are no elements in A which are less than or equal to k. Since we had assumed that A contains some element less than or equal to $k+1$, but nothing less than or equal to k, we conclude that $k+1 \in A$, and that $k+1$ is the smallest element of A.

Next, we show that the Well-ordering Principle implies mathematical induction. Suppose $A \subset \mathbb{N}$ is such that

- $1 \in A$;
- $x \in A$ implies $x + 1 \in A$;
- $A \neq \mathbb{N}$.

Let $B = \mathbb{N} - A$. By assumption B is not empty. By the Well-ordering Principle B has a smallest element b. Since $1 \in A$, $b \neq 1$, and as a result $b - 1 \in \mathbb{N}$. On the other hand, $b - 1 < b$, and as we had assumed that b is the smallest element of B, this means $b - 1 \notin B$. Consequently, $b - 1 \in A$, and this last statement implies that $(b-1)+1 \in A$, i.e., $b \in A$, a contradiction. □

2.2 Divisibility and prime factorization

Definition 2.4. For integers a, b with $b \neq 0$, we say b *divides* a if there is a $c \in \mathbb{Z}$ such that $a = bc$. The integer b is then called a *divisor* of a, and a is called a *multiple* of b. In this case, we write $b \mid a$. A natural number p is called *prime* if it has exactly four distinct divisors. For integers a, b, n, with $n \neq 0$, we write $a \equiv b \bmod n$, and say *a is congruent to b modulo n*, if $n \mid a - b$.

For example, $3 \mid (-6)$ as $-6 = 3 \cdot (-2)$. The number 5 is a prime number, since its divisors are $\pm 1, \pm 5$; 6 is not a prime as it is divisible by $\pm 1, \pm 2, \pm 3, \pm 6$, and 1 is not a prime as it only has two divisors ± 1. Finally, $13 \equiv 7 \bmod 3$ as $3 \mid 13 - 7 = 6$. Congruence modulo 0 is equality.

The following lemma is an easy exercise; see Exercise 2.1.

Lemma 2.5. *For an integer n,* congruence modulo n *is an equivalence relation.*

Definition 2.6. The equivalence classes of the congruence relation are called *congruence classes modulo n*. The congruence class of an integer a modulo a non-zero integer n is denoted by $[a]_n$. The set of congruence classes modulo n is denoted by $\mathbb{Z}/n\mathbb{Z}$.

Lemma 2.7. *The set $\mathbb{Z}/n\mathbb{Z}$ has a group structure defined by*

$$[a]_n + [b]_n := [a + b]_n.$$

Proof. The identity of the operation is given by $[0]_n$. The inverse of the element $[a]_n$ is $[-a]_n$. Associativity is immediate from the associativity of addition of the group $\mathbb{Z}$. □

Theorem 2.8 (Division Algorithm). *For integers a, b, with $b \neq 0$, there are unique integers q_0, r_0 with $0 \leq r_0 < |b|$, such that*

$$a = bq_0 + r_0.$$

If we allow negative values of r, we can choose q_0, r_0 such that

1. $-\frac{|b|+1}{2} \leq r_0 \leq \frac{|b|-1}{2}$, *if b is odd;*
2. $-\frac{|b|}{2} + 1 \leq r_0 \leq \frac{|b|}{2}$, *if b is even.*

Proof. By replacing q by $-q$ if necessarily, it suffices to prove the theorem for $b > 0$. If a, b, define

$$S = \{a - bq \mid q \in \mathbb{Z}, a - bq \in \mathbb{N}\}.$$

It is clear that $S \subset \mathbb{N}$. We claim that S is non-empty. To see this, we recognize two cases:

- If $a > 0$, then set $q = 0$. In this case $a - 0b = a > 0$, and $a \in S$;
- If $a < 0$ and $b > 0$, let $q = 2a$. We have $a - qb = a - 2ab = -a(2b - 1) > 0$. Again, $S \neq \varnothing$.

Since S is non-empty, Property 2.2 implies that S has a smallest element, call it x. By the definition of S, there is $q \in \mathbb{Z}$ such that $x = a - bq$. We now claim $x \leq b$. If $x = a - bq > b$, then $x - b = a - (b + 1)q > 0$. This means $x - b \in S$, and since $x - b < x$, this contradicts the choice of x as the smallest element of S.

Next, if the smallest element $x = b$, then $a - (q + 1)x = x - b = 0$, and we set $q_0 = q + 1$ and $r_0 = 0$. If $x < b$, then we set $q_0 = q$ and $r_0 = x$.

Now that we know the first part of the theorem, we can proceed to prove the second part. Suppose b is odd—the proof for the even case is similar. By the first part of the theorem we can write

$$a = bq_0 + r_0$$

with $0 \leq r_0 < |b|$. If $0 \leq r_0 \leq \frac{|b|-1}{2}$ we are done, so assume $\frac{|b|-1}{2} < r_0 < |b|$. We have

$$a = bq_0 + |b| + (r_0 - |b|).$$

Note that $bq_0 + |b|$ is a multiple of b. Next, since $\frac{|b|-1}{2} < r_0 < |b|$ we have

$$\frac{|b| - 1}{2} - |b| < r_0 - |b| < |b| - |b| = 0.$$

To finish the proof we need to verify that

$$\frac{|b| - 1}{2} - |b| \geq -\frac{|b| + 1}{2},$$

but this is clear. □

Note that with the notations of Theorem 2.8, $[a]_b = [r_0]_b$. This observation provides a convenient way to write down representatives for equivalence classes in $\mathbb{Z}/b\mathbb{Z}$. For example, suppose $b = 6$. When we divide an integer a by b, we will have a remainder 0, 1, 2, 3, 4, 5. Consequently, the set $\{0, 1, 2, 3, 4, 5\}$ will provide a set of representatives for $\mathbb{Z}/6\mathbb{Z}$.

Lemma 2.9. *For every non-zero integer n, $\#(\mathbb{Z}/n\mathbb{Z}) = |n|$.*

Proof. We define a map

$$res_n : \mathbb{Z}/n\mathbb{Z} \to \{0, 1, \cdots, |n| - 1\}.$$

The strategy of the proof is to show that the map res_n is a bijection. We define the function as follows. Let $u \in \mathbb{Z}/n\mathbb{Z}$. Let a be an integer such that $[a]_n = u$. Use Theorem 2.8 to write

$$a = qn + r$$

with $0 \leq r < |n|$. We define $res_n(u) = r$.

Since the definition of res_n involves a choice of the integer a, we need to show that $res_n(u)$ is independent of the choice of a. Suppose the integer b is such that $[b]_n = [a]_n = u$. The assumption on b implies that $a \equiv b \bmod n$, i.e., there is an integer k such that $b - a = kn$. If we use the fact that $a = qn + r$, we get $b = a + kn = qn + r + kn = (q + k)n + r$ with $0 \leq r < qn$. As a result, $res_n([b]_n) = r = res_n([a]_n)$.

We now show that res_n is a bijection. That it is a surjective map is obvious. In fact, for every r with $0 \leq r < n$, $res_n([r]_n) = r$. To see that it is injective, we suppose that $res_n(u) = res_n(u') = r$ with $u, u' \in \mathbb{Z}/n\mathbb{Z}$ and some r with the property that $0 \leq r < n$. Write $u = [a]_n$ and $u' = [b]_n$. It follows from the definition of res_n that $a = q_1 n + r$ and $b = q_2 n + r$ for integers q_1, q_2. As a result, $a - b = q_1 n - q_2 n = (q_1 - q_2)n$. Consequently, $n \mid a - b$, or $a \equiv b \bmod n$. This means $[a]_n = [b]_n$. □

Definition 2.10. Let n be an integer. By a *complete system of residues modulo n* we mean a collection of n integers $a_1, \ldots, a_n$ such that for each i, j with $1 \leq i, j \leq n$, we have $a_i \equiv a_j \bmod n$ if and only if $i = j$. Alternatively, a complete system of residues is a complete set of representatives for congruence classes modulo n.

The notion of the *greatest common divisor* described in the following definition is surprisingly important:

Definition 2.11. For integers a, b, the *greatest common divisor of* a, b, denoted $\gcd(a, b)$, is an integer g with the following properties:

- $g \mid a$ and $g \mid b$;
- If d is an integer such that $d \mid a$ and $d \mid b$, then $|d| \leq g$.

Integers a, b are called *coprime* if $\gcd(a, b) = 1$. We also define the *least common multiple of* the non-zero integers a, b, denoted by $\operatorname{lcm}(a, b)$ to be a positive integer l with the following properties:

- $a \mid l$ and $b \mid l$;
- If m is an integer such that $a \mid m$ and $b \mid m$, then $l \leq |m|$.

Basically, the greatest common divisor of integers a and b is precisely what the name suggests: the greatest, common, divisor of a and b, and similarly for the lcm. We similarly define the gcd and lcm of more than two numbers.

Theorem 2.12. *If a, b are integers, then there are integers x, y such that*

$$ax + by = \gcd(a, b).$$

Proof. The theorem is easy if either of a or b is zero. For example, if $a = 0$, then $\gcd(0, b) = b = 1 \times 0 + 1 \times b$. So we may assume that neither a nor b is zero. By changing the signs of x, y, if necessarily, we may assume $a, b > 0$. Define a set S by

$$S = \{ax + by \mid x, y \in \mathbb{Z}, ax + by \in \mathbb{N}\}.$$

Clearly $S \subset \mathbb{N}$ and $S \neq \varnothing$ as, in particular, $a, b \in S$. By Property 2.2, the set S has a smallest element g. By definition, there are integers x_0, y_0 such that $g = ax_0 + by_0$ and $g > 0$.

If d is a common divisor of a, b, then $d \mid ax_0 + by_0 = g$. Consequently, $\gcd(a, b) \mid g$.

Now we claim every element of S is divisible by g. Let $s = ax + by \in S$. Divide s by g, and use Theorem 2.8 to write

$$s = gq + r$$

for some $0 \leq r < g$. If $r = 0$, it follows that $g \mid s$ and we are done. Otherwise, we have

$$0 < r = s - gq = (ax + by) - (ax_0 + by_0)q = a(x - x_0q) + b(y - y_0q).$$

As a result, $r \in S$. Since $0 < r < g$, this last statement contradicts the assumption that g is the smallest element of S. Consequently, we have established the claim that every element of S is divisible by g. In particular, since $a, b \in S$, we see that $g \mid a$ and $g \mid b$, i.e., g is a common divisor of a, b. As a result, $g \leq \gcd(a, b)$. Since we

had already established that $\gcd(a, b) \mid g$, we conclude $g = \gcd(a, b)$. We have proved

$$\gcd(a, b) = ax_0 + by_0. \qquad \square$$

A consequence of this theorem is the following interesting result:

Corollary 2.13. *If a, b, d are integers such that $d \mid a$, $d \mid b$, then $d \mid \gcd(a, b)$.*

Proof. Since d is a divisor of both a and b, for all integers x, y we have $d \mid ax + by$. The result now follows from Theorem 2.12. $\square$

Clearly, one way to find the greatest common divisor of a and b is to write the list of all divisors of a and b, look for the common divisors, and find the greatest one. For example, if $a = 12$ and $b = 18$, we have

$$\text{Divisors of } a = \{\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 12\}$$

and

$$\text{Divisors of } b = \{\pm 1, \pm 2, \pm 3, \pm 6, \pm 9, \pm 18\}.$$

Next,

$$\text{Common divisors of } a \text{ and } b = \{\pm 1, \pm 2, \pm 3, \pm 6\}.$$

Finally,

$$\gcd(a, b) = 6.$$

Note that $6 = (+1) \cdot 18 + (-1) \cdot 12$ in accordance with Theorem 2.12.

This is, of course, inefficient, especially when dealing with large numbers. Euclid presented a clever procedure to compute the greatest common divisor of two integers without listing the divisors of the individual integers. This is known as the *Euclidean Algorithm*. The Euclidean Algorithm is based on the following lemma:

Lemma 2.14. *If $a, b \in \mathbb{N}$ with $a \mid b$, then $\gcd(a, b) = a$. If $a, b \in \mathbb{N}$ with $a > b$, then*

$$\gcd(a, b) = \gcd(a - b, b).$$

Proof. The first statement is easy. In fact, $\gcd(a, b) \leq a$ as the $\gcd(a, b)$ is a divisor of a. On the other hand, a is a common divisor of a and b, hence $a \leq \gcd(a, b)$. Combining these two observations shows that $\gcd(a, b) = a$. Now we prove the second statement by showing that the set of common divisors of a, b is equal to the set of common divisors of $a - b, b$. This statement implies that the greatest elements of the sets are the same, proving the lemma. To see the equality of the two sets, suppose d is a common divisor of a, b. Then $d \mid a$, $d \mid b$, and consequently $d \mid a - b$, i.e., d is a common divisor of b and $a - b$. Hence, the set of common divisors of a, b is a subset of the set of common divisors of b and $a - b$. The reverse inclusion is proved similarly. $\square$

As an example, we compute $\gcd(18, 12)$. We have

$$\gcd(18, 12) = \gcd(18 - 12, 12) = \gcd(6, 12) = 6,$$

by applying Lemma 2.14. To see a slightly more interesting example, we examine $\gcd(57, 12)$. We have

$$\gcd(57, 12) = \gcd(57 - 12, 12) = \gcd(45, 12) = \gcd(33, 12)$$

$$= \gcd(21, 12) = \gcd(9, 12) = \gcd(12, 9) = \gcd(12 - 9, 9) = \gcd(3, 9) = 3.$$

In the first stage, we needed to subtract 12 from 57 four times. In effect, what we have done is that we have replaced 57 by the remainder of its division by 12. In practice, we do the following: In order to compute $\gcd(a, b)$ with $a > b$, we write $a = bq + r$ with $0 \leq r < b$; if $r = 0$, then $\gcd(a, b) = b$; otherwise, $\gcd(a, b) = \gcd(b, r)$. Since $a > b > r$, we have replaced the pair (a, b) with the "smaller" pair (b, r) with the same gcd. Let us formulate this procedure as a lemma:

Lemma 2.15 (Euclidean Algorithm). *The following procedure computes the* gcd *of a pair of natural numbers* (a, b) *with* $a > b$*:*

1. *The pair* (a, b) *is given with* $a > b$*;*
2. *Let* r *be the remainder of the division of* a *by* b*;*
3. *If* $r = 0$*,* b *is the* gcd *and we are done;*
4. *If* $r > 0$*, replace* (a, b) *by* (b, r)*, and go back to (1).*

At the time of this writing, we do not know how to find the prime factors of a large integer n quickly. In contrast, the Euclidean Algorithm is incredibly fast. In fact, by Theorem 12 of [46, Ch. I, §3], originally a theorem of Lamé from 1844, the number of divisions needed is at most five times the number of digits in the decimal expansion of the smaller number b.

The Euclidean Algorithm allows us to make Theorem 2.12 computationally effective. We will illustrate the idea in the following example:

Example 2.16. It is easy to see that $\gcd(57, 12) = 3$. We wish to find integers x, y such that

$$57x + 12y = 3.$$

We write

$$57 = 4 \times 12 + 9;$$

$$12 = 1 \times 9 + 3.$$

Now we write

$$3 = 12 - 9 = 12 - (57 - 4 \times 12) = 12 - 57 + 4 \times 12 = 5 \times 12 - 57,$$

giving $x = -1$ and $y = 5$. We will see more examples of this procedure in the exercises.

A consequence of Theorem 2.12 is the following important theorem:

Theorem 2.17. *If $a \mid bc$ and $\gcd(a, b) = 1$, then $a \mid c$.*

Proof. Since $\gcd(a, b) = 1$, there are integers x, y such that $ax + by = 1$. Multiplying the equality by c gives $c = axc + bcy$. Both terms on the right-hand side of this equation are divisible by a: The term axc is clearly divisible by a, and bcy is divisible by a by assumption. This means c is divisible by a and we are done. □

This theorem implies the following result of Euclid (Elements, Proposition 30, Book VII):

Corollary 2.18 (Euclid's First Theorem). *Let p be a prime number, and $p \mid ab$ for integers a, b. Then either $p \mid a$ or $p \mid b$.*

Proof. Suppose $p \nmid a$. We claim that $\gcd(a, p) = 1$. In fact, if $d = \gcd(a, p)$, then $d \mid p$. This means that either $d = 1$ or $d = p$. We cannot have $d = p$, because then $p = d \mid a$ which is a contradiction. Hence, $d = 1$, and the result follows from Theorem 2.17. □

Euclid's Lemma is the main ingredient in the proof of the uniqueness assertion of the following foundational result:

Theorem 2.19 (Fundamental Theorem of Arithmetic). *Every natural number is a product of prime numbers in an essentially unique fashion.*

In the statement of the theorem, *essentially unique* means up to reordering of the terms. For example, we can write

$$12 = 3 \cdot 2 \cdot 2 = 2 \cdot 3 \cdot 2 = 2 \cdot 2 \cdot 3.$$

Proof. We will prove the existence using induction. Since 1 is the *empty* product of prime numbers, the theorem is true for 1. Now suppose n is a natural number, and suppose we know the existence of a prime factorization for every natural number smaller than n. If n is prime, there is nothing to prove. If n is not prime, then it has a non-trivial divisor y such that $1 < y < n$. Clearly, $1 < n/y < n$. By the induction assumption, $y = p_1 \cdots p_r$ and $n/y = q_1 \cdots q_s$ for primes $p_1, \ldots, p_r$ and $q_1, \ldots, q_s$. Then,

$$n = y \cdot \frac{n}{y} = p_1 \cdots p_r \cdot q_1 \cdots q_s.$$

This gives the existence of a prime factorization.

We now prove the uniqueness. Suppose we have a natural number n which has two different prime factorizations:

$$P_1 \cdots P_k = Q_1 \cdots Q_l.$$

The sets of primes $\{P_1, \ldots, P_k\}$ and $\{Q_1, \ldots, Q_l\}$ may have some common elements. If necessary we simplify the common elements from the sides to obtain an equality of the form

$$P_1 \cdots P_u = Q_1 \cdots Q_v, \tag{2.1}$$

with the sides not having any common factors. Now, we have

$$P_1 \mid Q_1 \cdots Q_v.$$

An easy application of Euclid's First Theorem, Corollary 2.18, says that there is an i such that

$$P_1 \mid Q_i.$$

But since P_1 and Q_i are prime numbers, this divisibility implies that $P_1 = Q_i$, contradicting the assumption that the sides of Equation (2.1) have no common elements. □

It is convenient to write the prime factorization of a number as a product of prime powers. For example, instead of $12 = 2 \cdot 3 \cdot 2$, we usually write $12 = 2^2 \cdot 3$. We denote the prime factorization of a typical natural number n in the form

$$p_1^{\alpha_1} \cdots p_r^{\alpha_r},$$

or similar expression. In such expressions, even when we do not explicitly mention it, we assume that the prime numbers $p_1, \ldots, p_r$ are distinct. In this case we write $p_i^{\alpha_i} || n$, meaning $p_i^{\alpha_i} | n$ but $p_i^{\alpha_i+1} \nmid n$, and call α_i the *multiplicity* of p_i in n. It is sometimes convenient to allow the exponents α_i to be equal to zero. For example, if

$$n = p_1^{\alpha_1} \cdots p_r^{\alpha_r},$$

then every divisor of n can be written in the form

$$m = p_1^{\beta_1} \cdots p_r^{\beta_r},$$

where for each i, $0 \leq \beta_i \leq \alpha_i$.

The Fundamental Theorem of Arithmetic has many applications. Here we list three consequences. We leave the proofs to the reader; see Exercise 2.4 and Exercise 2.5.

Proposition 2.20. *Let $m = \prod_i p_i^{r_i}$ and $n = \prod_i p_i^{s_i}$. Then*

$$\gcd(m, n) = \prod_i p_i^{\min(r_i, s_i)},$$

and

$$\operatorname{lcm}(m, n) = \prod_i p_i^{\max(r_i, s_i)}.$$

Furthermore,

$$\gcd(m, n) \cdot \operatorname{lcm}(m, n) = mn.$$

The following proposition is used a few times throughout the book:

Proposition 2.21. *Suppose a, b are natural numbers such that $\gcd(a, b) = 1$. If $ab = m^k$ for natural numbers m and k, then $a = m_1^k$ and $b = m_2^k$ for natural numbers m_1, m_2 such that $m_1 m_2 = m$.*

Corollary 2.22. *If $n \in \mathbb{N}$ is not a perfect kth power, there is no rational number γ such that $n = \gamma^k$.*

2.3 The Chinese Remainder Theorem

Theorem 2.12 is a statement about the solvability of the equation

$$ax + by = \gcd(a, b)$$

in integers x, y. More generally, one can ask about the solvability of a general linear Diophantine equation

$$ax + by = c$$

in integers x, y. It is not hard to see that this equation is solvable if and only if $\gcd(a, b) \mid c$. For example if $\gcd(a, b) = 1$, then every equation $ax + by = c$ is solvable. The following is a useful fact:

Theorem 2.23. *Suppose a, b are coprime integers, and let $x_0, y_0 \in \mathbb{Z}$ be such that $ax_0 + by_0 = 1$. Then if $x, y \in \mathbb{Z}$ satisfy $ax + by = 1$, there is $h \in \mathbb{Z}$ such that*

$$x = x_0 + bh, \quad y = y_0 - ah.$$

In general, if the equation $ax+by = c$ is solvable, then since $\gcd(a, b) \mid ax+by$, we see that $\gcd(a, b) \mid c$. Conversely, if $\gcd(a, b) \mid c$, we can write $c = c' \cdot \gcd(a, b)$. By Theorem 2.12 we know that there are integers x_0, y_0 such that $ax_0 + by_0 = \gcd(a, b)$. Multiplying by c' gives $a(x_0c') + b(y_0c') = \gcd(a, b)c' = c$, and as a result $x = x_0c'$ and $y = y_0c'$ are numbers that satisfy $ax + by = c$.

Formulated in terms of congruence equations, this is equivalent to saying that the equation

$$ax \equiv c \mod b \tag{2.2}$$

is solvable if and only if $\gcd(a, b) \mid c$. In particular if $\gcd(a, b) = 1$, the equation is solvable for every c. Back in the general case of Equation (2.2), since $\gcd(a, b) \mid c$, the equation is equivalent to

$$\frac{a}{\gcd(a, b)}x \equiv \frac{c}{\gcd(a, b)} \mod \frac{b}{\gcd(a, b)}. \tag{2.3}$$

Now

$$\gcd\left(\frac{a}{\gcd(a, b)}, \frac{b}{\gcd(a, b)}\right) = 1,$$

and as a result Equation (2.2) is solvable with solution

$$x \equiv \left(\frac{a}{\gcd(a, b)}\right)^{-1} \frac{c}{\gcd(a, b)} \mod \frac{b}{\gcd(a, b)}.$$

So every equation of the form (2.2), if solvable, has a solution of the form

$$x \equiv k \mod m$$

for some $m \mid b$.

For example, the equation $4x \equiv 3 \bmod 6$ is not solvable as $2 = \gcd(4, 6) \nmid 3$. On the other hand, the equation $4x \equiv 2 \bmod 6$ is solvable as $2 = \gcd(4, 6) \mid 2$. To solve the equation $4x \equiv 2 \bmod 6$, we divide by 2 to get $2x \equiv 1 \bmod 3$, which has the solution $x \equiv 2 \bmod 3$.

One can also ask about the solvability of systems of equations

$$\begin{cases} a_1 x \equiv c_1 \mod b_1, \\ a_2 x \equiv c_2 \mod b_2. \end{cases}$$

Obviously we need each of the equations to be solvable, so our previous considerations apply. In particular the solvability of this system reduces to the solvability of a system of the form

$$\begin{cases} x \equiv k_1 \mod m_1, \\ x \equiv k_2 \mod m_2. \end{cases} \tag{2.4}$$

It is not hard to see, Exercise 2.22, that this system is solvable if and only if

$$\gcd(m_1, m_2) \mid k_1 - k_2.$$

If x_1, x_2 are solutions of the system (2.4), then $x_1 \equiv x_2 \bmod [m_1, m_2]$.

For a system consisting of more than two equations the exact solvability conditions are fairly painful to state. However, there is a useful special case with many applications:

Theorem 2.24 (The Chinese Remainder Theorem). *Suppose $m_1, \ldots, m_n$ are integers such that for all i, j with $i \neq j$,*

$$\gcd(m_i, m_j) = 1.$$

Then for every string of integers $a_1, \ldots, a_n$ the system of equations

$$\begin{cases} x \equiv a_1 \mod m_1, \\ \quad \cdots \\ x \equiv a_n \mod m_n, \end{cases}$$

has a solution. If x_1, x_2 are solutions of the system, then

$$x_1 \equiv x_2 \mod m_1 \cdots m_n.$$

Example 2.25. Suppose we wish to find all x such that

$$\begin{cases} x \equiv 1 \mod 5; \\ x \equiv 2 \mod 7; \\ x \equiv 3 \mod 9. \end{cases}$$

Every x satisfying the first equation is of the form $1 + 5k$. Insert this expression in the second equation to obtain

$$1 + 5k \equiv 2 \mod 7.$$

This is the same as saying $5k \equiv 1 \mod 7$, which after multiplying by 3 gives $k \equiv 3 \mod 7$, i.e., $k = 3+7l$ for some l. This means, $x = 1+5k = 1+5(3+7l) = 16+35l$. Now we use the third equation to obtain

$$16 + 35l \equiv 3 \mod 9.$$

Since $16 \equiv -2$ and $35 \equiv -1 \mod 9$, we get $-2 - l \equiv 3 \mod 9$, from which it follows $l \equiv 4 \mod 9$. Write $l = 4 + 9r$ for some $r \in \mathbb{Z}$. Going back to x, we have $x = 16 + 35l = 16 + 35(4 + 9r) = 156 + 315r$. Consequently, in order for x to satisfying the system of congruences it is necessary and sufficient that

$$x \equiv 156 \mod 315.$$

2.4 Euler's Theorem

Next, we discuss a beautiful theorem of Fermat:

Theorem 2.26 (Fermat's Little Theorem). *If p is prime, for all integers n, $p \mid n^p - n$.*

First we consider $p = 2$. We know that n is even if and only if n^2 is even. For this reason $n^2 - n$ is always divisible by 2, establishing the theorem for $p = 2$. So we assume that p is an odd prime. In this case it is clear that if the theorem is true for n, it will also be true for $-n$. It suffices to prove the theorem for n a natural number. We proceed by induction. The theorem is trivially true for $n = 0, 1$. Now suppose the theorem is true for n. We wish to prove it is true for $n+1$. By the Binomial Theorem, Theorem A.4, we have

$$(n + 1)^p - (n + 1) = (n^p - n) + \sum_{k=1}^{p-1} \binom{p}{k} n^k.$$

Since by our induction hypothesis, $p \mid n^p - n$, the theorem follows from the following lemma:

Lemma 2.27. *For each $0 < k < p$,*

$$p \mid \binom{p}{k}.$$

Proof. We have

$$\binom{p}{k} = \frac{p!}{k!(p-k)!} = \frac{p \cdot (p-1)!}{k!(p-k)!}.$$

Since $\binom{p}{k}$ is an integer, this means $k!(p-k)! \mid p.(p-1)!$; but since $\gcd(p, k!(p-k)!) = 1$, Theorem 2.17 implies $k!(p-k)! \mid (p-1)!$. Write $(p-1)! = k!(p-k)! \cdot A$ for an integer A. Then

$$\binom{p}{k} = \frac{p \cdot (p-1)!}{k!(p-k)!} = p \cdot A.$$

The lemma is now obvious. □

We will record one more lemma that will be used in the proof of Theorem 6.8 in Chapter 7.

Lemma 2.28. *Let p be a prime number, and $x_1, \ldots, x_n$ some indeterminates. Then all of the coefficients of the multivariable polynomial*

$$(x_1 + \cdots + x_n)^p - x_1^p - \ldots x_n^p$$

are integers that are multiples of p.

We now describe Euler's generalization of Fermat's Little Theorem. The following proposition is an easy consequence of Theorem 2.12:

Proposition 2.29. *If a and n with $\gcd(a, n) = 1$, then there exists an integer b such that $ab \equiv 1 \bmod n$.*

Proof. Since $\gcd(a, n) = 1$, Theorem 2.12 implies that there are integers b and c such that $ab + cn = 1$. This means $n \mid ab - 1$, i.e., $ab \equiv 1 \bmod n$.

For example if $a = 3$ and $n = 7$, then we may take $b = 5$, as in that case $3 \times 5 \equiv 1 \bmod 7$. The congruence class of the b in the proposition is usually denoted by a^{-1} when there is no confusion about the modulus n. This means that the set of coprime to n congruence classes forms a group under multiplication modulo n. We denote this group by $(\mathbb{Z}/n\mathbb{Z})^\times$.

Definition 2.30. Let $n \in \mathbb{N}$. By a *reduced system of residues modulo n* we mean a set of representatives for $(\mathbb{Z}/n\mathbb{Z})^\times$. For a natural number n, we define the *Euler totient function*, or Euler's ϕ-function, by $\phi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$.

For every complete system of residues $a_1, \ldots, a_n$ modulo n, the set

$$\{a_i \mid \gcd(a_i, n) = 1\} \tag{2.5}$$

is a reduced system of residues. It is clear that every reduced system of residues modulo n has the same number of elements, $\phi(n)$. Furthermore, if $a_1, \ldots, a_{\phi(n)}$ is a set of distinct residue classes modulo n such that for each i we have $\gcd(a_i, n) = 1$, then the set $a_1, \ldots, a_{\phi(n)}$ is a reduced system of residues modulo n. Note that

$$\phi(n) = \#\{1 \le a \le n \mid \gcd(a, n) = 1\}. \tag{2.6}$$

If, for example, $n = 12$, then the numbers a with $1 \le a \le 12$ which are coprime to 12 are 1, 5, 7, 11, and consequently, $\phi(12) = 4$.

Theorem 2.31 (Euler). *Let n be a natural number. For all a with* $\gcd(a, n) = 1$ *the equation*

$$a^{\phi(n)} \equiv 1 \mod n$$

holds.

In particular when $n = p$ is a prime number, we have $\phi(p) = p - 1$, and we recover Fermat's Little Theorem, Theorem 2.26.

Proof. Suppose $a_1, \ldots, a_{\phi(n)}$ is a reduced system of residues modulo n. Since $\gcd(a, n) = 1$, the set of numbers

$$aa_1, \ldots, aa_{\phi(n)}$$

is another reduced system of residues modulo n. In fact, for each $1 \leq i \leq \phi(n)$, $\gcd(aa_i, n) = 1$. Furthermore, as $\gcd(a, n) = 1$, $aa_i \equiv aa_j \mod n$ for $1 \leq i, j \leq \phi(n)$ implies $a_i \equiv a_j \mod n$, which means $i = j$. Next, since

$$a_1, \ldots, a_{\phi(n)}$$

and

$$aa_1, \ldots, aa_{\phi(n)}$$

are both reduced systems of residues, we must have

$$\prod_{i=1}^{\phi(n)} a_i \equiv \prod_{i=1}^{\phi(n)} aa_i \mod n.$$

Rearranging terms gives

$$\prod_{i=1}^{\phi(n)} a_i \equiv a^{\phi(n)} \prod_{i=1}^{\phi(n)} a_i \mod n.$$

Since the a_i's are coprime to n, their product is coprime to n as well. Simplifying $\prod_i a_i$ gives the result. □

The function $\phi(n)$ is explicitly computable. It is easy to see that for each prime p and $\alpha \geq 1$ we have

$$\phi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha(1 - \frac{1}{p}).$$

In fact,

$$\begin{aligned}\phi(p^\alpha) &= p^\alpha - \#\{1 \leq a \leq p^\alpha \mid \gcd(a, p^\alpha) \neq 1\} \\ &= p^\alpha - \#\{1 \leq a \leq p^\alpha \mid p|a\} \\ &= p^\alpha - p^{\alpha-1}.\end{aligned}$$

The totient function is famously *multiplicative*:

Theorem 2.32. *For all natural numbers m, n with $\gcd(m, n) = 1$, the identity*

$$\phi(mn) = \phi(m)\phi(n)$$

holds.

Proof. We prove this theorem by constructing a reduced system of residues modulo mn. For $r, s \in \mathbb{Z}$, set

$$f(r, s) = rn + sm.$$

Our first claim is that the set

$$R = \{f(r, s) \mid 1 \leq r \leq m, 1 \leq s \leq n\}$$

is a complete system of residues modulo mn. Clearly, we have mn pairs (r, s) as above. We just need to show that for distinct pairs (r, s), the elements $f(r, s)$ are distinct modulo mn. Suppose, for $1 \leq r_1, r_2 \leq m$ and $1 \leq s_1, s_2 \leq n$, we have

$$r_1 n + s_1 m \equiv r_2 n + s_2 m \mod mn.$$

Considering this congruence modulo n gives

$$s_1 m \equiv s_2 m \mod n,$$

which, since $\gcd(m, n) = 1$, implies

$$s_1 \equiv s_2 \mod n.$$

Since $1 \leq s_1, s_2 \leq n$, this gives $s_1 = s_2$. Similarly, we conclude $r_1 = r_2$, and our claim is proved.

Next, we claim that in order for $\gcd(f(r, s), mn) = 1$, it is necessary and sufficient that $\gcd(r, m) = 1$ and $\gcd(s, n) = 1$. In fact, since $\gcd(m, n) = 1$, we have

$$\gcd(f(r, s), mn) = \gcd(rn + sm, m) \gcd(rn + sm, n).$$

Next,

$$\gcd(rn + sm, m) = \gcd(rn, m) = \gcd(r, m).$$

Similarly,

$$\gcd(rn + sm, n) = \gcd(s, n).$$

Consequently,

$$\gcd(f(r, s), mn) = \gcd(r, m) \cdot \gcd(s, n),$$

from which the second claim is immediate. The number of all pairs (r, s) such that $\gcd(r, m) = 1$ and $\gcd(s, n) = 1$ is clearly $\phi(m)\phi(n)$, and the theorem is proved. □

It then follows that for each natural number n with prime factorization $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ we have

$$\phi(n) = \prod_{i=1}^{k} \phi(p_i^{\alpha_i}) = \prod_{i=1}^{k} p_i^{\alpha_i} \left(1 - \frac{1}{p_i}\right) = n \prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right).$$

We will record this computation as a theorem:

Theorem 2.33. *For every natural number n,*

$$\phi(n) = n \prod_{p|n} (1 - \frac{1}{p}).$$

This theorem means that in order to compute the value of $\phi(n)$ we just need to know the prime factors of n, and not the prime factorization. For example, since the prime factors of 12 are 2 and 3, we have

$$\phi(12) = 12 \cdot \left(1 - \frac{1}{2}\right) \cdot \left(1 - \frac{1}{3}\right) = 12 \times \frac{1}{2} \times \frac{2}{3} = 4.$$

Theorem 2.33 has an interesting statistical interpretation. Suppose we have a number n with prime factors $p_1, p_2, \ldots, p_k$. The quotient $\phi(n)/n$ is the probability of choosing a random number a in the set $\{1, \ldots, n\}$ subject to $\gcd(a, n) = 1$. Now, a number a satisfies $\gcd(a, n) = 1$ if and only if for each i, $p_i \nmid a$. The probability of a randomly chosen number to be divisible by p_i is $1/p_i$, and the probability that a randomly chosen number is coprime to p_i is $1 - 1/p_i$. If we pretend that coprimality to distinct primes are independent events, we see that the probability that a number is coprime to $p_1, \ldots, p_k$ is

$$\prod_{i=1}^{k} \left(1 - \frac{1}{p_i}\right),$$

which by Theorem 2.33 is precisely $\phi(n)/n$.

The function ϕ viewed as a function $\mathbb{N} \to \mathbb{R}$ has many surprising properties. Here is an example:

Theorem 2.34. *For all natural numbers n,*

$$\sum_{d|n} \phi(d) = n.$$

Proof. By Theorem A.2 there are precisely n distinct complex numbers z such that $z^n = 1$, and they can be expressed as

$$e^{\frac{2\pi i k}{n}}, \quad k = 0, \ldots, n-1.$$

For a complex number z with $z^n = 1$, we define $o(z)$ to be the smallest positive integer k such that $z^k = 1$. We claim $o(z) \mid n$. If not, by Theorem 2.8 there is an integer q and $0 < r < o(z)$ such that $n = qo(z) + r$. Then

$$1 = z^n = z^{qo(z)+r} = (z^{o(z)})^q z^r = z^r,$$

contradicting the definition of $o(z)$. Next,

$$n = \#\{z \in \mathbb{C} \mid z^n = 1\} = \sum_{d|n} \#\{z \in \mathbb{C} \mid z^n = 1, o(z) = d\}. \tag{2.7}$$

Our next step is to determine $\#\{z \in \mathbb{C} \mid z^n = 1, o(z) = d\}$. In order to do this, we pick $0 \leq k \leq n-1$ and determine $o(e^{\frac{2\pi ik}{n}})$. Suppose for $l > 0$ we have

$$\left(e^{\frac{2\pi ik}{n}}\right)^l = 1.$$

This is equivalent to saying

$$e^{\frac{2\pi ikl}{n}} = 1.$$

Consequently, $n \mid kl$. Dividing by $\gcd(n,k)$ gives

$$\frac{n}{\gcd(n,k)} \mid \frac{k}{\gcd(n,k)} \cdot l.$$

Since

$$\gcd\left(\frac{n}{\gcd(n,k)}, \frac{k}{\gcd(n,k)}\right) = 1,$$

Theorem 2.17 implies that

$$\frac{n}{\gcd(n,k)} \mid l.$$

This statement combined with $l > 0$ implies

$$l \geq \frac{n}{\gcd(n,k)}.$$

In particular,

$$o(e^{\frac{2\pi ik}{n}}) \geq \frac{n}{\gcd(n,k)}.$$

We claim that equality holds. To see this, we note

$$\left(e^{\frac{2\pi ik}{n}}\right)^{\frac{n}{\gcd(n,k)}} = e^{\frac{2\pi ik}{n} \cdot \frac{n}{\gcd(n,k)}} = e^{2\pi i \cdot \frac{k}{\gcd(n,k)}} = 1,$$

as $\frac{k}{\gcd(n,k)}$ is an integer. Hence, we have

$$o(e^{\frac{2\pi ik}{n}}) = \frac{n}{\gcd(n,k)}.$$

Now we can go back to determining $\#\{z \in \mathbb{C} \mid z^n = 1, o(z) = d\}$. If $o(e^{\frac{2\pi ik}{n}}) = d$, then we have $\frac{n}{\gcd(n,k)} = d$. It follows, $\gcd(n,k) = \frac{n}{d}$. In particular, $\frac{n}{d} \mid k$. Write $k = \frac{n}{d} \cdot k'$. Note $1 \leq k' \leq d$. We have

$$\frac{n}{d} = \gcd(n,k) = \gcd\left(\frac{n}{d} \cdot d, \frac{n}{d} \cdot k'\right) = \frac{n}{d} \cdot \gcd(d,k').$$

Hence $\gcd(d,k') = 1$. This means,

$$\#\{z \in \mathbb{C} \mid z^n = 1, o(z) = d\} = \#\{1 \leq k' \leq d \mid \gcd(d,k') = 1\} = \phi(d).$$

Combining this identity with (2.7) gives the theorem. □

2.5 Polynomials modulo a prime

We often speak of *polynomials modulo* p, with p a prime number. By this we mean a polynomial $f(x) \in \mathbb{Z}[x]$ where the coefficients and values are considered modulo p. This is of course nothing but a polynomial in the variable x with coefficients in the finite field $\mathbb{Z}/p\mathbb{Z}$, but for the purposes of this monograph we can prove the results we need completely elementarily using the methods presented in this chapter.

Throughout this discussion we fix a prime number p. Let $f(x) = \sum_{j=0}^{n} a_j x^j$, with $a_j \in \mathbb{Z}$, be a polynomial. We say f is a *non-zero polynomial* modulo p, if there is a j with $a_j \not\equiv 0 \bmod p$; we say f is of *degree* n if $a_n \not\equiv 0 \bmod p$. We call an integer k *a root of* $f(x)$ *modulo* p if $f(k) \equiv 0 \bmod p$. We call the roots k, l distinct if $k \not\equiv l \mod p$. For example, if $p = 3$, the polynomial $f(x) = x^5 + 2$ is of degree 5 and has a root $k = 1$ modulo 3. One easily checks that $l = 4$ is another root of $f(x)$ modulo 3, but 1 and 4 are not distinct modulo 3, as $4 \equiv 1 \bmod 3$.

Remark 2.35. Note that these notions depend on the choice of the prime p. For example, if $f(x) = 3x^4 + 2x + 5$, then $f(x)$ is of degree 4 if $p \neq 3$, but of degree 1 for $p = 3$. Also, $f(2) = 57 = 3 \times 19$, so $k = 2$ is a root of $f(x)$ modulo 3 and 19, but not otherwise.

Our goal in this section is to prove the following useful statement:

Theorem 2.36. *Let $f(x)$ be a polynomial of degree n modulo a prime p. Then $f(x)$ has at most n distinct roots modulo p.*

Our proof of this theorem relies on the following lemma the statement of which the reader should compare with Theorem 2.8:

Lemma 2.37. *Suppose $f(x)$ and $g(x)$ are polynomials with integer coefficients, and suppose $g(x)$ is a monic polynomial. Then there are unique polynomials $q(x)$ and $r(x)$ with integral coefficients such that*

$$f(x) = q(x)g(x) + r(x),$$

and either $r(x) = 0$ or $0 \leq \deg r(x) < \deg g(x)$.

Proof. We will prove the lemma by induction on $\deg f$. If $\deg f < \deg g$, then there is nothing to prove, as we can simply set $q(x) = 0$ and $r(x) = f(x)$. Now suppose $\deg f \geq \deg g$, and write $f(x) = \sum_{j=0}^{n} a_j x^j$ and $g(x) = x^m + \sum_{l=0}^{m-1} b_l x^l$ with $a_n \neq 0$. Then $\deg(f(x) - a_n x^{n-m} g(x)) < \deg f$. By induction, there are polynomials $q'(x), r'(x)$ with integer coefficients such that either $r'(x) = 0$ or $\deg r'(x) < \deg g(x)$, and with the property that

$$f(x) - a_n x^{n-m} g(x) = q'(x)g(x) + r'(x).$$

This equation implies $f(x) = (q'(x) + a_n x^{n-m})g(x) + r'(x)$. Setting $q(x) = q'(x) + a_n x^{n-m} \in \mathbb{Z}[x]$ and $r(x) = r'(x)$ gives the result. For uniqueness see Exercise 2.31. □

For example, if we divide the polynomial $f(x) = 3x^3 + 2$ by the polynomial $g(x) = x^2 + 2$, we get $q(x) = 3x$ and $r(x) = -6x + 2$, and $\deg r(x) < \deg g(x)$.

Proof of Theorem 2.36. We prove this theorem by induction on the degree of the polynomial f. If $f(x)$ has no roots modulo p, then there is nothing to prove. So suppose we have a root k. Use Lemma 2.37 to write

$$f(x) = (x - k)q(x) + r(x).$$

The lemma says that either $r(x) = 0$ or $\deg r(x) < \deg(x - k) = 1$. This means that either $r(x) = 0$ or $\deg r(x) = 0$, i.e., $r(x)$ is a constant c which may be zero. In any case, we write

$$f(x) = (x - k)q(x) + c.$$

Insert $x = k$ in this expression to obtain $f(k) = (k - k)q(k) + c = c$. So we obtain

$$f(x) = (x - k)q(x) + f(k).$$

Since $f(k) \equiv 0 \bmod p$, we have

$$f(x) \equiv (x - k)q(x) \mod p.$$

Consequently, the roots of $f(x)$ modulo p consist of k plus whatever root modulo p that $q(x)$ may have. Since $\deg q(x) = \deg f(x) - 1$, by induction, $q(x)$ has at most $\deg f(x) - 1$ roots. The result is now immediate. □

Remark 2.38. In general, if F is a field, and $f(x) \in F[x]$ a non-zero polynomial, then $f(x) = 0$ has at most $\deg f(x)$ roots in F.

Theorem 2.36 has numerous applications. The following example is a particularly well-known application of this theorem.

Example 2.39. Fix a prime p. By Theorem 2.26, for all $n \in \mathbb{Z}$ we have

$$n^p - n \equiv 0 \mod p.$$

This means that if we set

$$f(x) = x^p - x,$$

then every integer n is a root of $f(x)$ modulo p. As a result, the elements of a complete system of residues modulo p, e.g., $S = \{0, 1, \ldots, p - 1\}$, are going to be the distinct roots of $f(x)$. On the other hand, we have a polynomial

$$g(x) = x(x - 1)(\cdots)(x - p + 1) = x^p + \text{ terms of lower degree.}$$

with the elements of S as its roots. Now consider the polynomial

$$h(x) = f(x) - g(x).$$

It is clear that $\deg h(x) < p$ as the x^p terms from the polynomial $f(x)$ and $g(x)$ cancel each other out. Now, every element of S is a root of $h(x)$ modulo p. But this

would mean that the polynomial h which is of degree less than p has p roots which contradicts Theorem 2.36, unless $h(x) \equiv 0 \mod p$. Consequently,

$$x^p - x \equiv x(x-1)(\cdots)(x-p+1) \mod p.$$

We may cancel out an x from the congruence to obtain the identity

$$x^{p-1} - 1 \equiv (x-1)(\cdots)(x-p+1) \mod p.$$

If we put in $x = p$ in this identity we obtain the following statement known as Wilson's Theorem:

$$(p-1)! \equiv -1 \mod p. \tag{2.8}$$

2.6 Digit expansions

It is common practice to express real numbers in terms of powers of 10. We call such expressions *decimal expansions*. For example, when we write $x = 347$, what we mean is that x is equal to the following expression

$$3 \times 10^2 + 4 \times 10^1 + 7 \times 10^0.$$

In this expression the numbers 3, 4, 7 are called the *digits* of x. The digits are always integers larger than or equal to 0 and less than or equal to 9. If we have a non-integral real number, then we use a decimal point to separate the integer part from the fractional part. For instance, when we write $x = 23.6923$ we mean

$$x = 2 \times 10^2 + 3 \times 10^0 + 6 \times 10^{-1} + 9 \times 10^{-2} + 2 \times 10^{-3} + 3 \times 10^{-4}.$$

We wish to generalize this notion. Suppose $g > 1$ is a natural number. In this section we discuss *base g expansions* of real numbers. We will show that every positive real number x is representable in the form

$$\sum_{k \in \mathbb{Z}, k < N} a_k \cdot g^k$$

with $N \in \mathbb{Z}$ and a_k's integers satisfying $0 \leq a_k < g$. Once we establish this, we write

$$x = (a_{N-1} \cdots a_1 a_0 . a_{-1} a_{-2} a_{-3} \cdots)_g,$$

if $N > 0$, and

$$x = (0 . a_{-1} a_{-2} a_{-3} \cdots)_g$$

if $N = 0$, and

$$x = (0.0 \cdots 0 a_{N-1} a_{N-2} \cdots)_g$$

if $N < 0$, where the number of zeros between the decimal point and a_{N-1} is $-N$. We will also determine the extent to which this representation is unique. Throughout

the remainder of this section we will use Exercise 2.47 without explicit mention numerous times.

Suppose $x \in \mathbb{R}$ and $x > 0$. We write $x = n + \xi$ with $n \in \mathbb{N} \cup \{0\}$ and $0 \leq \xi < 1$. Here, $n = [x]$ and $\xi = \{x\}$. Our first step is to construct the base g expansion of n.

If $n = 0$, then we define the base g expansion of n to be 0. So we assume $n \geq 1$. By the Well-ordering Principle, Property 2.2, there is a smallest natural number N such that $n < g^N$. This means that

$$g^{N-1} \leq n < g^N,$$

as otherwise $n < g^{N-1}$ which would contradict the choice of N. By Theorem 2.8 there are integers q and n' such that

$$n = q \cdot g^{N-1} + n'$$

with $0 \leq n' < g^{N-1}$. We claim $0 \leq q < g$. In fact, if $q > g$, then

$$n = q \cdot g^{N-1} + n' > g \cdot g^{N-1} = g^N,$$

contradicting the choice of N; if $q \leq -1$, then

$$n = q \cdot g^{N-1} + n' < -g^{N-1} + n' < 0.$$

Now that we know $0 \leq q < g$, we denote it by a_{N-1}. We have

$$n = a_{N-1} \cdot g^{N-1} + n'$$

with $0 \leq n' < g^{N-1}$. By repeating this process we obtain the representation

$$n = a_{N-1} \cdot g^{N-1} + \cdots + a_1 \cdot g^1 + a_0 \tag{2.9}$$

with each a_i satisfying $0 \leq a_i < g$. The expression on the right-hand side of (2.9) is the *base g expansion of n*, and the a_i's are called the *digits* of n.

Now we show that the base g expansion of a natural number is unique. Suppose a natural number n has two different base g expansions:

$$n = a_{N-1} \cdot g^{N-1} + \cdots + a_1 \cdot g^1 + a_0 = b_{M-1} \cdot g^{M-1} + \cdots + b_1 \cdot g^1 + b_0, \tag{2.10}$$

with $M, N > 0$ and $0 \leq a_i, b_j < g$, and let's assume $a_{N-1} \neq 0, b_{M-1} \neq 0$. First we show $M = N$. Suppose $M > N$. We observe

$$b_{M-1} \cdot g^{M-1} + \cdots + b_1 \cdot g^1 + b_0 \geq g^{M-1}.$$

Next,

$$a_{N-1} \cdot g^{N-1} + \cdots + a_1 \cdot g^1 + a_0 \leq (g-1)g^{N-1} + (g-1)g^{N-2} + \cdots + (g-1)g + (g-1)$$

$$= (g-1)(g^{N-1} + \cdots + g + 1) = (g-1)\frac{g^N - 1}{g-1} = g^N - 1 < g^{M-1}.$$

So on the one hand $n \geq g^{M-1}$ and on the other $n < g^{M-1}$. This is a contradiction, showing that M cannot be larger than N. Similarly, it follows that $N > M$ is impossible as well. As a result $M = N$.

With the equality $M = N$ at hand, Equation (2.10) can be rewritten as

$$a_{N-1} \cdot g^{N-1} + \cdots + a_1 \cdot g^1 + a_0 = b_{N-1} \cdot g^{N-1} + \cdots + b_1 \cdot g^1 + b_0.$$

We will show that for each i, $a_i = b_i$. We will first show that $a_{N-1} = b_{N-1}$. After this has been established, the rest of the argument is an easy induction. Suppose $a_{N-1} \neq b_{N-1}$. Then we have

$$\begin{aligned} 0 &= \left|(a_{N-1} \cdot g^{N-1} + \cdots + a_1 \cdot g^1 + a_0) - (b_{N-1} \cdot g^{N-1} + \cdots + b_1 \cdot g^1 + b_0)\right| \\ &= \left|(a_{N-1} - b_{N-1})g^{N-1} + (a_{N-2} - b_{N-2})g^{N-2} + \cdots + (a_1 - b_1)g + (a_0 - b_0)\right| \\ &\geq \left|(a_{N-1} - b_{N-1})g^{N-1}\right| - \left|(a_{N-2} - b_{N-2})g^{N-2} + \cdots + (a_1 - b_1)g + (a_0 - b_0)\right|, \end{aligned}$$

upon using the following version of *the triangle inequality*: For all real numbers x, y, $|x + y| \geq |x| - |y|$. Since $a_{N-1} \neq b_{N-1}$,

$$\left|(a_{N-1} - b_{N-1})g^{N-1}\right| \geq g^{N-1}.$$

Also, for each i, $|a_i - b_i| \leq g - 1$. This inequality implies

$$\begin{aligned} &\left|(a_{N-2} - b_{N-2})g^{N-2} + \cdots + (a_1 - b_1)g + (a_0 - b_0)\right| \\ &\leq \left|(a_{N-2} - b_{N-2})g^{N-2}\right| + \cdots + |(a_1 - b_1)g| + |(a_0 - b_0)| \\ &\leq (g-1)g^{N-2} + \cdots + (g-1)g + (g-1) = g^{N-1} - 1, \end{aligned}$$

after using the triangle inequality in the following form: For all $x_1, x_2, \ldots, x_k \in \mathbb{R}$, we have $|x_1 + x_2 + \cdots + x_k| \leq |x_1| + |x_2| + \cdots + |x_k|$. Putting everything together, we have

$$\begin{aligned} 0 \geq \left|(a_{N-1} - b_{N-1})g^{N-1}\right| - &\left|(a_{N-2} - b_{N-2})g^{N-2} + \cdots + (a_1 - b_1)g + (a_0 - b_0)\right| \\ &\geq g^{N-1} - (g^{N-1} - 1) = 1. \end{aligned}$$

This is a contradiction, showing that $a_{N-1} = b_{N-1}$. We have proved the following lemma:

Lemma 2.40. *Let $g \in \mathbb{N}$ and $g > 1$. Then every natural number n can be written in a unique way as a sum*

$$n = a_{N-1} \cdot g^{N-1} + \cdots + a_1 \cdot g^1 + a_0$$

with $N \in \mathbb{N}$ and $a_i \in \mathbb{N} \cup \{0\}$ satisfying $0 \leq a_i < g$.

The integers a_j are called the *digits* of n, and we write

$$n = (a_{N-1} \ldots a_1)_g.$$

Now we construct the base g expansion of ξ, the fractional part of the real number x. Set

$$a_{-1} = [g\xi].$$

Next for each $k > 1$, set

$$a_{-k} = \left[g^k \left(\xi - \sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} \right) \right] = [g^k \xi] - g^k \left(\sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} \right).$$

For example,

$$a_{-2} = \left[g^2 \left(\xi - \frac{a_{-1}}{g} \right) \right] = [g^2 \xi] - g \cdot a_{-1},$$

and

$$a_{-3} = \left[g^3 \left(\xi - \frac{a_{-1}}{g} - \frac{a_{-2}}{g^2} \right) \right] = [g^3 \xi] - g^2 \cdot a_{-1} - g \cdot a_{-2}.$$

Now we claim that for each $k > 0$,

$$0 \le \xi - \sum_{j=1}^{k} \frac{a_{-j}}{g^j} < \frac{1}{g^k}. \tag{2.11}$$

If $k = 1$, then by the definition of the integer part we have

$$0 \le g\xi - [g\xi] < 1.$$

Since $a_{-1} = [g\xi]$, this gives $0 < g\xi - a_{-1} < 1$, from which upon dividing by g our inequality follows. For $k > 1$, we have

$$0 \le g^k \left(\xi - \sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} \right) - \left[g^k \left(\xi - \sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} \right) \right] < 1.$$

By definition this means

$$0 \le g^k \left(\xi - \sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} \right) - a_{-k} < 1.$$

Dividing by g^k gives

$$0 \le \xi - \sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} - \frac{a_{-k}}{g^k} < \frac{1}{g^k},$$

and this is the inequality (2.11).

Since $0 < \xi < 1$, $0 < g\xi < g$, and as a result $0 \le [g\xi] < g$. This means $0 \le a_{-1} < g$. If $k > 1$, (2.11) implies

$$0 \leq g^k \left(\xi - \sum_{j=1}^{k-1} \frac{a_{-j}}{g^j} \right) < g,$$

which gives

$$0 \leq a_{-k} < g.$$

Lemma 2.41. *With a_j's defined as above,*

$$\xi = \sum_{j=1}^{\infty} \frac{a_{-j}}{g^j}. \tag{2.12}$$

Proof. Once we note that $g^{-k} \to 0$ as k gets large, this is a consequence of Equation (2.11). □

As before we call the integers a_{-j}'s the *digits* of ξ, and we write

$$\xi = (0.a_{-1}a_{-2}a_{-3}\dots)_g$$

and call it the *base g expansion* of ξ.

On the other hand we can consider expressions of the form

$$\sum_{j=1}^{\infty} \frac{a_{-j}}{g^j} \tag{2.13}$$

with a_j's integers satisfying $0 \leq a_{-j} < g$ and ask whether they correspond to base g expansions of real numbers. First a lemma:

Lemma 2.42. *Every expression of the form (2.13) is convergent.*

Proof. In order to see this, set

$$s_N = \sum_{j=1}^{N} \frac{a_{-j}}{g^j}.$$

By the definition of convergence, for $\varepsilon > 0$, we need to show there is N_0 such that if $M, N > N_0$, then

$$|s_N - s_M| < \varepsilon.$$

Without loss of generality suppose $N > M$. Then,

$$|s_N - s_M| = \sum_{j=M+1}^{N} \frac{a_{-j}}{g^j} \leq \sum_{j=M+1}^{N} \frac{g-1}{g^j} = \frac{g-1}{g^{M+1}} \sum_{k=0}^{N-M-1} \frac{1}{g^k}$$

$$= \frac{g-1}{g^{M+1}} \cdot \frac{1 - \frac{1}{g^{M+N}}}{1 - \frac{1}{g}} = \frac{1}{g^M} \cdot \frac{g^{M+N} - 1}{g^{M+N}} < \frac{1}{g^M}.$$

So given $\varepsilon > 0$, we pick N_0 such that

$$\frac{1}{g^{N_0}} < \varepsilon.$$

Once this is done, the above computation shows that as soon as $N > M > N_0$, then

$$|s_N - s_M| < \frac{1}{g^M} < \frac{1}{g^{N_0}} < \varepsilon,$$

establishing the convergence. □

Now we ask whether distinct series of the sort considered in Equation (2.13) can give the same real number. Suppose we have an identity

$$\sum_{j=1}^{\infty} \frac{a_{-j}}{g^j} = \sum_{j=1}^{\infty} \frac{b_{-j}}{g^j},$$

where each side is a series of the type considered above: For each j, a_{-j}, b_{-j} are integers satisfying $0 \leq a_{-j}, b_{-j} < g$. Let N be the smallest natural number such that $a_{-N} \neq b_{-N}$. Then we have

$$0 = \sum_{j=1}^{\infty} \frac{a_{-j}}{g^j} - \sum_{j=1}^{\infty} \frac{b_{-j}}{g^j} = \left| \sum_{j=1}^{\infty} \frac{a_{-j}}{g^j} - \sum_{j=1}^{\infty} \frac{b_{-j}}{g^j} \right|$$

$$= \left| \sum_{j=N}^{\infty} \frac{a_{-j} - b_{-j}}{g^j} \right| \geq \left| \frac{a_{-N} - b_{-N}}{g^N} \right| - \sum_{j=N+1}^{\infty} \left| \frac{a_{-j} - b_{-j}}{g^j} \right|$$

$$\geq \frac{1}{g^N} - \sum_{j=N+1}^{\infty} \frac{g-1}{g^j} = 0,$$

using an easy computation involving geometric series. As a result all the inequalities appearing here should be equalities. This means that either $a_{-N} - b_{-N} = 1$ and for each $j > N$, $a_{-j} = 0, b_{-j} = g - 1$, or $a_{-N} - b_{-N} = -1$, and for each $j > N$, $a_{-j} = g-1, b_{-j} = 0$. What this means, for example, is that if we have a sequence of integers b_{-j}, with $0 \leq b_{-j} < g$ such that for some N, and for all $j > N, b_{-j} = g-1$, then we can define a real number ξ by setting

$$\xi = \sum_{j=1}^{\infty} \frac{b_{-j}}{g^j}.$$

Now if we write the base g digit expansion of ξ according to Lemma 2.41 we obtain

$$\xi = \frac{b_{-1}}{g} + \cdots + \frac{b_{-N+1}}{g^{N-1}} + \frac{1 + b_{-N}}{g^N}. \tag{2.14}$$

We call such a base g expansion a *finite expansion*. We say an expansion of the form

$$\sum_{j=1}^{\infty} \frac{b_{-j}}{g^j}$$

to be *unacceptable* if there is M such that for all $j \geq M$, $b_{-j} = g - 1$. We say an expansion is *acceptable* if it is not unacceptable.

It is clear that the number ξ with expansion as in Equation (2.14) can be written as

$$\xi = \frac{r}{g^N}$$

for some natural number r. By canceling out every common factor between r and g^N we arrive at a fraction of the form A/B where all the prime factors of B are prime factors of g. Conversely, suppose we have a fraction of the form $\xi = A/B$ with

$$B = \prod_{p|g, p \text{ prime}} p^{e_p}$$

with integers $e_p \geq 0$. Let $M = \max_p e_p$, the largest number among the e_p's. Let

$$C = \prod_{p|g, p \text{ prime}} p^{M-e_p}.$$

Then $BC = g^M$. We have

$$\xi = \frac{A}{B} = \frac{AC}{BC} = \frac{r}{g^M}$$

with $r = AC$. Now we write the base g expansion of r using Lemma 2.40 in the form

$$r = \sum_{k=0}^{N-1} a_k \cdot g^k$$

for some $N \in \mathbb{N}$, $0 \leq a_k < g$. We then have

$$\xi = \frac{\sum_{k=0}^{N-1} a_k \cdot g^k}{g^M} = \sum_{k=0}^{N-1} a_k \cdot g^{k-M}.$$

We summarize this discussion as the following proposition:

Proposition 2.43. *An expression of the form*

$$\sum_{j=1}^{\infty} \frac{b_{-j}}{g^j}$$

is the base g expansion of some real number $0 \leq \xi < 1$ if and only if it is acceptable. The base g expansion of a real number ξ is finite if and only if it is a rational number expressible in the form A/B with B a divisor of g^M for some M.

Putting everything together, we have the following theorem:

Theorem 2.44. *Let $g > 1$ be a natural number. Every positive real number can be written as*

$$\sum_{k\in\mathbb{Z},k<N} a_k \cdot g^k$$

with $N \in \mathbb{N}$, $a_k \in \mathbb{N} \cup \{0\}$, $0 \leq a_k < g$, subject to the additional requirement that

$$\sum_{k\in\mathbb{Z},k<0} a_k \cdot g^k$$

be acceptable.

2.7 Digit expansions of rational numbers

In Proposition 2.43 we determined the base g expansions of rational numbers A/B with B certain special numbers. In this section we determine what base g expansions of arbitrary rational numbers look like.

The reader will probably remember from elementary school that decimal expansions of rational numbers are eventually periodic, in the sense that there will be blocks of digits that will repeat exactly. For example:

$$\frac{7}{15} = 0.46666666\ldots;$$

$$\frac{2}{7} = 0.285714285714285714\ldots;$$

$$\frac{7}{12} = 0.583333333\ldots;$$

$$\frac{1}{19} = 0.052631578947368421052631578947368421052631578947368421\ldots.$$

In the first example, the repeating block is the single digit 6; in the second one, it is 285714; in the third one, 3; and in the last one, 052631578947368421. The common practice is to draw a line above the repeating block so as to save space and avoid confusion, e.g.,

$$\frac{7}{17} = 0.4\overline{6}; \quad \frac{2}{7} = \overline{285714}; \quad \frac{7}{12} = 0.58\overline{3}; \quad \frac{1}{19} = 0.\overline{052631578947368421}.$$

We will see that similar results hold for base g expansions of rational numbers for arbitrary natural numbers $g > 1$. We say a base g expansion is *repeating* if from some point onward, the sequence of digits is the back to back repetitions of some fixed finite sequence of numbers. The examples we gave above are all repeating expansions for the base 10. In general, a repeating base g expansion will look like this:

$$(a_{N-1}\ldots a_1.a_{-1}a_{-2}\ldots a_{-k}\overline{b_1b_2\ldots b_t})_g, \tag{2.15}$$

where as before the line on top of $b_1 b_2 \dots b_t$ means that this is the repeating sequence of digits. We call the sequence $b_1 \dots b_k$ the *repeating block*, and the number k, the *period*. A base g expansion of the form $(a_{N-1} \dots a_1.\overline{b_1 b_2 \dots b_k})_g$ is called *purely periodic*.

Our goal is to prove the following theorem:

Theorem 2.45. *Let $g > 1$ be a natural number. A positive real number is rational if and only if its base g expansion is repeating.*

Proof. Note that a finite base g expansion is repeating: The repeating sequence of numbers is simply 0. We already saw in Proposition 2.43 that finite base g expansions give rational numbers.

Let $g > 1$ be a natural number. Our first step is to show that repeating base g expansions give rational numbers. Suppose we have a repeating base g expansion as in Equation (2.15):

$$\begin{aligned} x &= (a_{N-1} \dots a_1.a_{-1}a_{-2} \dots a_{-k}\overline{b_1 b_2 \dots b_t})_g \\ &= (a_{N-1} \dots a_1.a_{-1}a_{-2} \dots a_{-k})_g + \frac{(0.\overline{b_1 b_2 \dots b_t})_g}{g^{k+1}}. \end{aligned}$$

By Proposition 2.43, or just by direct inspection, the number $(a_{N-1} \dots a_1.a_{-1}a_{-2} \dots a_{-k})_g$ is rational. So in order to show that x is rational, we just need to show that

$$\gamma := (0.\overline{b_1 b_2 \dots b_t})_g$$

is a rational number. In order to see this we observe

$$\begin{aligned} \gamma &= \sum_{j=0}^{\infty} \left(\frac{b_1}{g^{1+jk}} + \frac{b_2}{g^{2+jk}} + \dots + \frac{b_k}{g^{k+jk}} \right) \\ &= \left(\frac{b_1}{g^1} + \frac{b_2}{g^2} + \dots + \frac{b_k}{g^k} \right) \sum_{j=0}^{\infty} \frac{1}{g^{jk}} \\ &= \frac{b_1 \cdot g^{k-1} + b_2 \cdot g^{k-2} + \dots + b_k}{g^k(1 - g^{-k})}, \end{aligned}$$

after using Exercise 2.47. We conclude that

$$\gamma = \frac{b_1 \cdot g^{k-1} + b_2 \cdot g^{k-2} + \dots + b_k}{g^k - 1}, \tag{2.16}$$

clearly showing that γ is a rational number. We have shown that every repeating base g expansion gives a rational number.

Next we show that the base g expansion of every positive rational number is repeating. Suppose we have a rational number

$$x = \frac{A}{B}$$

with $A, B \in \mathbb{N}$, and $\gcd(A, B) = 1$. The starting point of the argument is to write

$$B = B_1 \cdot B_2$$

with B_1 the largest divisor of B which is coprime to g. This means that every prime factor of B_2 is prime factor of g. By an argument similar to the one used in the paragraph preceding Proposition 2.43 there is an integer C and a natural number M such that $B_2 C = g^M$. We then have

$$x = \frac{A}{B} = \frac{AC}{BC} = \frac{AC}{B_1 B_2 C} = \frac{AC}{B_1 g^M}.$$

Note that if we show the base g expansion of AC/B_1 is repeating, then we will be done, as dividing by g^M only introduces a shift in the base g expansion. So without loss of generality, we may assume that

$$x = A/B$$

with $A, B \in \mathbb{N}$, $\gcd(A, B) = 1$, $\gcd(B, g) = 1$. By Theorem 2.8 we can write

$$A = qB + r$$

with $0 \leq r < B$. This means

$$x = q + \frac{r}{B}.$$

If $r = 0$ there is nothing to prove, so suppose $r \neq 0$. It suffices to show that the base g expansion of r/B is repeating. The key to the argument is the expression we found in Equation (2.16). Suppose there is an integer D such that $BD = g^k - 1$ for some $k \in \mathbb{N}$. Then we have

$$\frac{r}{B} = \frac{rD}{BD} = \frac{rD}{g^k - 1}.$$

Now since $rD < g^k - 1$, reversing the steps of the first part of the proof shows that the base g expansion of r/B is repeating. So in order to finish the proof we just need to prove the following assertion: If B with $\gcd(B, g) = 1$, then there is a $k \in \mathbb{N}$ such that $B \mid g^k - 1$, i.e., $g^k \equiv 1 \mod B$. By Theorem 2.31 $k = \phi(B)$ works and we are done. □

2.8 Primitive roots

In the proof of Theorem 2.45 we observed that if g, B with $\gcd(g, B) = 1$, then for each $0 < r < B$, the fraction r/B has a purely periodic base g expansion with period $\phi(B)$. In general the fraction r/B may have a smaller period. For example, let's consider the fraction $1/7$. The base 2 expansion of the fraction $1/7$ can be computed as follows:

$$\frac{1}{7} = \frac{1}{2^3 - 1} = \frac{1}{2^3} \cdot \frac{1}{1 - 2^{-3}} = \frac{1}{2^3} \cdot \sum_{k=0}^{\infty} \frac{1}{2^{3k}} = \sum_{k=0}^{\infty} \frac{1}{2^{3k+3}}.$$

From this computation it follows that

$$\frac{1}{7} = (0.001001001001\ldots)_2 = (0.\overline{001})_2.$$

The period is 3, which is half of $\phi(7) = 6$. Now we compute the base 3 expansion of $1/7$:

$$\frac{1}{7} = \frac{1}{3^6 - 1} = \frac{1}{3^6} \cdot \frac{1}{1 - 3^{-6}} = \frac{1}{3^6} \cdot \sum_{k=0}^{\infty} \frac{1}{3^{6k}} = \sum_{k=0}^{\infty} \frac{1}{3^{6k+6}}.$$

Consequently,

$$\frac{1}{7} = (0.000001000001000001\ldots)_3 = (0.\overline{000001})_3,$$

and in this case the period is 6. So depending on the base g, sometimes the period of the base g expansion of $1/7$ is $\phi(7) = 6$, and sometimes it is not. In fact, it follows from the proof of Theorem 2.45 that the minimal period of the base g expansion of $1/n$, if $\gcd(g, n) = 1$, is the smallest positive integer k such that $g^k \equiv 1 \mod n$. We make the following definition.

Definition 2.46. For a natural number n, and an integer a, with $\gcd(a, n) = 1$, the *order of a modulo n*, denoted by $o_n(a)$, is the smallest positive integer k such that

$$a^k \equiv 1 \mod n.$$

Note that by Theorem 2.31, $o_n(a) \leq \phi(n)$. Also, the congruence classes of the elements

$$a^j, \quad 1 \leq j \leq o_n(a)$$

are distinct modulo n.

Lemma 2.47. *If for some integer k, $a^k \equiv 1 \mod n$, then $o_n(a) \mid k$. In particular, $o_n(a) \mid \phi(n)$.*

Proof. Write $k = qo_n(a) + r$ with $0 \leq r < o_n(a)$. We have,

$$1 \equiv a^k \equiv a^{qo_n(a)+r} \equiv \left(a^{o_n(a)}\right)^q a^r \equiv (1)^q a^r \equiv a^r \mod n.$$

Consequently, $a^r \equiv 1 \mod n$. Since $0 \leq r < o_n(a)$, this last equation implies $r = 0$. The last assertion follows from Theorem 2.31. □

Definition 2.48. A number g is called a *primitive root modulo n* if $o_n(g) = \phi(n)$.

In terms of fractions, this means that the base g expansion of $1/n$ is purely periodic of period $\phi(n)$, the largest possible value. The existence of a primitive root is equivalent to the cyclicity of the Abelian group $(\mathbb{Z}/n\mathbb{Z})^\times$. Note that primitive roots may not

exist. For example, if $n = 8$, then $\phi(n) = 4$. However, for all odd numbers a, $a^2 \equiv 1 \bmod 8$. In fact, $1^2 \equiv 1, 3^2 = 9 \equiv 1, 5^2 = 25 \equiv 1$, and $7^2 = 49 \equiv 1 \bmod 8$. In contrast, if $n = 7$, then $3^1 \equiv 3 \bmod 7$, $3^2 \equiv 2 \bmod 7$, $3^3 \equiv 6 \bmod 7$, $3^4 \equiv 4 \bmod 7$, $3^5 \equiv 5 \bmod 7$, and $3^6 \equiv 1 \bmod 7$, implying that 3 is a primitive root modulo 7.

The following theorem provides an extremely important class of situations where we know primitive roots exist.

Theorem 2.49. *If p is a prime, there is a primitive root modulo p.*

Proof. If $p = 2$, then the result is trivial. So let's assume p is odd. Let $a_1, \ldots, a_{p-1}$ be a reduced system of residues modulo p. Since by Lemma 2.47 for each j, $o_p(a_j) \mid p - 1$, we have

$$p - 1 = \sum_{d \mid p-1} \#\{1 \leq j \leq p - 1 \mid o_p(a_j) = d\}. \tag{2.17}$$

In our case, Theorem 2.34 says

$$p - 1 = \sum_{d \mid p-1} \phi(d). \tag{2.18}$$

Our strategy is to show that for each $d \mid p - 1$,

$$\#\{1 \leq j \leq p - 1 \mid o_p(a_j) = d\} = \phi(d). \tag{2.19}$$

Once this is established, letting $d = p - 1$ gives

$$\#\{1 \leq j \leq p - 1 \mid o_p(a_j) = p - 1\} = \phi(p - 1) \neq 0.$$

This means there are primitive roots modulo p, and in fact $\phi(p - 1)$ of them.

We now proceed to prove (2.19). Our first step is to show that for $d \mid p - 1$, if $\#\{1 \leq j \leq p - 1 \mid o_p(a_j) = d\} \neq 0$, then it is equal to $\phi(d)$. So, let us assume that this quantity is non-zero and pick a congruence class a modulo p such that $o_p(a) = d$. Since the congruence classes of the d elements

$$a^j, \quad 1 \leq j \leq d$$

are distinct and all satisfy the equation $x^d \equiv 1 \bmod p$, Theorem 2.36 implies that these are all the solutions of the equation.

Now we determine which of these elements a^j have the property $o_p(a^j) = d$. In order to do this, for an integer k, with $1 \leq k \leq d$, let us determine $o_p(a^k)$. If for a positive integer l, $(a^k)^l \equiv 1 \bmod p$, we get $a^{kl} \equiv 1 \bmod p$. Lemma 2.47 implies that $o_p(a) \mid kl$, or $d \mid kl$. This implies

$$\frac{d}{\gcd(d, k)} \mid \frac{k}{\gcd(d, k)} \cdot l.$$

Since

$$\gcd\left(\frac{d}{\gcd(d, k)}, \frac{k}{\gcd(d, k)}\right) = 1,$$

Theorem 2.17 implies $\frac{d}{\gcd(d,k)} \mid l$. We conclude

$$l \geq \frac{d}{\gcd(d,k)}.$$

In particular,

$$o_p(a^k) \geq \frac{d}{\gcd(d,k)}.$$

As in the proof of Theorem 2.34, we claim that equality holds. It suffices to check that

$$(a^k)^{\frac{d}{\gcd(d,k)}} \equiv 1 \mod p.$$

But this is immediate, as

$$(a^k)^{\frac{d}{\gcd(d,k)}} \equiv (a^d)^{\frac{k}{\gcd(d,k)}} \equiv 1 \mod p.$$

We have used the fact that $a^d \equiv 1 \bmod p$ and $k/\gcd(d,k) \in \mathbb{N}$. Now that we have established

$$o_p(a^k) = \frac{d}{\gcd(d,k)},$$

we determine under what conditions on k, $o_p(a^k) = d$. In order for this to happen we need to have

$$\frac{d}{\gcd(d,k)} = d,$$

or, what is the same, $\gcd(d,k) = 1$. Consequently, if $1 \leq k \leq d$ with $\gcd(d,k) = 1$, $o_p(a^k) = d$. As a result, if $o_p(a) = d$, then

$$\{a^k \mid 1 \leq k \leq d, \gcd(d,k) = 1\}$$

is the set of elements whose congruence classes have order d modulo p. Since the latter set has $\phi(d)$ elements, we conclude that if

$$\#\{1 \leq j \leq p-1 \mid o_p(a^j) = d\} \neq 0,$$

then

$$\#\{1 \leq j \leq p-1 \mid o_p(a^j) = d\} = \phi(d).$$

As a result, for each $d \mid p-1$,

$$\phi(d) - \#\{1 \leq j \leq p-1 \mid o_p(a^j) = d\} \geq 0.$$

Summing up over all $d \mid p-1$ gives

$$\sum_{d|p-1} \left(\phi(d) - \#\{1 \leq j \leq p-1 \mid o_p(a^j) = d\}\right) = 0,$$

after using (2.17) and (2.18). Since each term of the sum is nonnegative this means every term has to be zero, establishing (2.19). The proof of the theorem is complete. □

Remark 2.50. If p is a small prime number, then it is easy to check whether a number g is a primitive root. For example, one can check easily, by direct computation, that $g = 2$ is a primitive root modulo 11. For large primes it is in general difficult to decide if a natural number g is a primitive root modulo p. Later in Lemma 2.57 we present a criterion to decide whether a number g is a primitive root modulo a prime p. This criterion, unfortunately, requires the knowledge of the prime factors of $p-1$.

Next we use the above theorem to determine all numbers n for which there is a primitive root modulo n.

Theorem 2.51. *There is a primitive root modulo n if and only if $n = 1, 2, 4, p^\alpha, 2p^\alpha$, for an odd prime p.*

We present the proof of this theorem as a series of lemmas.

Lemma 2.52. *Suppose n can be written as mk, with $\gcd(m,k) = 1$ and $m, k > 2$. Then there are no primitive roots modulo n. In particular, if there is a primitive root modulo n, then $n = 2^\alpha, p^\alpha, 2p^\alpha$ for some odd prime p.*

Proof. Let a be an integer such that $\gcd(a,n) = 1$. Then $\gcd(a,m) = \gcd(a,k) = 1$. By Theorem 2.31, we have $a^{\phi(m)} \equiv 1 \mod m$ and $a^{\phi(k)} \equiv 1 \mod k$. Since $\phi(m) \mid \operatorname{lcm}(\phi(m), \phi(k))$,

$$a^{\operatorname{lcm}(\phi(m),\phi(k))} \equiv 1 \mod m.$$

Similarly,

$$a^{\operatorname{lcm}(\phi(m),\phi(k))} \equiv 1 \mod k.$$

By the uniqueness assertion of Theorem 2.24 we have

$$a^{\operatorname{lcm}(\phi(m),\phi(k))} \equiv 1 \mod mk. \tag{2.20}$$

Next, we observe that $\operatorname{lcm}(\phi(m), \phi(k)) < \phi(mk)$. Indeed,

$$\operatorname{lcm}(\phi(m), \phi(k)) = \frac{\phi(m)\phi(k)}{\gcd(\phi(m), \phi(k))} = \frac{\phi(mk)}{\gcd(\phi(m), \phi(k))},$$

after using Proposition 2.20 and Theorem 2.32. Since $m, k > 2$, Exercise 2.39 shows that $\phi(m)$ and $\phi(k)$ are both even, and consequently, $\gcd(\phi(m), \phi(k))$ is a non-zero even number, hence $\operatorname{lcm}(\phi(m), \phi(k)) < \phi(mk)$. Equation (2.20) now shows that there is an integer $0 < u < \phi(mk)$ such that for all a with $\gcd(a, mk) = 1$ we have $a^u \equiv 1 \mod mk$. This proves the lemma. □

It is clear that if $n = 1, 2$ then 1 is a primitive root modulo n. Also, if $n = 4$ then there is a primitive root, namely $g = 3$. Now we show that there are no primitive roots for higher powers of 2.

Lemma 2.53. *If $n = 2^\alpha$ with $\alpha > 2$, then there are no primitive roots modulo n.*

Proof. We have already seen that for all odd numbers a, $a^2 \equiv 1 \mod 8$. Since $\phi(8) = 4$, this means that for all a with $\gcd(a, 8) = 1$ we have

$$a^{\frac{\phi(2^3)}{2}} \equiv 1 \mod 2^3.$$

Our goal is to show that for all $\alpha > 2$, and for all a with $\gcd(a, 2^\alpha) = 1$, we have

$$a^{\frac{\phi(2^\alpha)}{2}} \equiv 1 \mod 2^\alpha. \tag{2.21}$$

Note that this identity proves the lemma. Since $\phi(2^\alpha) = 2^\alpha(1 - 1/2) = 2^{\alpha-1}$, (2.21) is equivalent to saying

$$a^{2^{\alpha-2}} \equiv 1 \mod 2^\alpha. \tag{2.22}$$

We will prove this assertion via mathematical induction. We already checked the validity of the claim for $\alpha = 3$. Now suppose we know (2.21) for α. This means there is an integer k such that

$$a^{2^{\alpha-2}} = 1 + k2^\alpha.$$

Next,

$$a^{2^{\alpha-1}} = \left(a^{2^{\alpha-2}}\right)^2 = (1+k2^\alpha)^2 = 1+2\cdot k2^\alpha+2^{2\alpha} = 1+k2^{\alpha+1}+2^{2\alpha} \equiv 1 \mod 2^{\alpha+1},$$

proving (2.22). The lemma has been proved. □

Remark 2.54. Compare the above proof with the proof of Lemma 8.5.

With these lemmas in place, we just need to prove the existence of primitive roots for $n = p^\alpha, 2p^\alpha$ for p an odd prime number. The key input is Theorem 2.49. First we prove the existence of primitive roots for the powers of an odd prime.

Lemma 2.55. *If p is an odd prime and $\alpha \in \mathbb{N}$, there is a primitive root modulo p^α.*

Proof. By Theorem 2.49 we know the result for $\alpha = 1$. Let g be a primitive root modulo p. We will show that either g or $g + p$ is a primitive root modulo p^2. We know that $o_{p^2}(g) \mid \phi(p^2) = p(p-1)$. On the other hand, since $g^{o_{p^2}(g)} \equiv 1 \mod p^2$, we have $g^{o_{p^2}(g)} \equiv 1 \mod p$. Consequently, $p - 1 = o_p(g) \mid o_{p^2}(g)$. This means that $o_{p^2}(g) \mid p(p-1)$ and $p - 1 \mid o_{p^2}(g)$. Consequently, there are two possibilities for $o_{p^2}(g)$: Either $o_{p^2}(g) = p(p-1) = \phi(p^2)$ in which case we have already found a primitive root modulo p^2, or $o_{p^2}(g) = p - 1$. Suppose we are in this latter situation. Since $g+p \equiv g \mod p$, $g+p$, too, is a primitive root modulo p, and again $o_{p^2}(g+p)$ is either $p - 1$ or $p(p-1)$. We will show that $o_{p^2}(g + p) \neq p - 1$. In order to see this we compute $(g + p)^{p-1}$ by using the Binomial Theorem (Theorem A.4) and we will show that it is not congruent to 1 modulo p^2. By Theorem A.4 we have

$$(g + p)^{p-1} = \sum_{k=0}^{p-1} \binom{p-1}{k} g^{p-1-k} p^k.$$

Now we examine this identity modulo p^2, noting that if $k \geq 2$, $p^k \equiv 0 \bmod p^2$. We have

$$(g+p)^{p-1} \equiv g^{p-1} + (p-1)g^{p-2}p \equiv 1 + (p-1)pg^{p-2} \mod p^2.$$

This expression is not congruent to 1 modulo p^2, since otherwise,

$$1 + (p-1)pg^{p-2} \equiv 1 \mod p^2$$

implies $p^2 \mid (p-1)pg^{p-2}$, or $p \mid (p-1)g^{p-2}$ which is impossible. Consequently, $o_{p^2}(g+p) = p(p-1) = \phi(p^2)$.

Now suppose that for $\alpha \geq 2$ we have a primitive root g modulo p^α. We will show that g is also a primitive root modulo $p^{\alpha+1}$. As before, $o_{p^{\alpha+1}}(g) \mid \phi(p^{\alpha+1}) = p^\alpha(p-1)$ and $p^{\alpha-1}(p-1) = \phi(p^\alpha) \mid o_{p^{\alpha+1}}(g)$. Again, there are two possibilities for $o_{p^{\alpha+1}}(g)$: Either it is equal to $\phi(p^{\alpha+1})$ in which case we are done, or it is equal to $\phi(p^\alpha)$. To reach a contradiction, let us assume

$$o_{p^{\alpha+1}}(g) = \phi(p^\alpha) = p^{\alpha-1}(p-1).$$

In particular,

$$g^{p^{\alpha-1}(p-1)} \equiv 1 \mod p^{\alpha+1}. \tag{2.23}$$

Let m be the largest nonnegative integer such that $p^m \mid g^{p^{\alpha-2}(p-1)} - 1$, so that

$$g^{p^{\alpha-2}(p-1)} = 1 + up^m \tag{2.24}$$

for some integer u with $\gcd(u, p) = 1$. Note that this is indeed a sensible definition as $\alpha \geq 2$. Furthermore, since by Theorem 2.26, $g^{p-1} \equiv 1 \bmod p$, $m \geq 1$. Next,

$$g^{p^{\alpha-1}(p-1)} = (g^{p^{\alpha-2}(p-1)})^p = (1+up^m)^p.$$

Applying Theorem A.4 gives

$$g^{p^{\alpha-1}(p-1)} = 1 + p \cdot up^m + \sum_{k=2}^{p} \binom{p}{k} (up^m)^k = 1 + u'p^{m+1}$$

with u' an integer satisfying $(u', p) = 1$. Going back to (2.23) we obtain

$$1 + u'p^{m+1} \equiv 1 \mod p^{\alpha+1}.$$

It then follows that $p^{m+1} \equiv 0 \bmod p^{\alpha+1}$, or what is the same, $m \geq \alpha$. Equation (2.24) now shows

$$g^{p^{\alpha-2}(p-1)} \equiv 1 \mod p^\alpha.$$

This is a contradiction as g was assumed to be a primitive root modulo p^α, and $p^{\alpha-2}(p-1) < p^{\alpha-1}(p-1) = \phi(p^\alpha)$. This contradiction shows that $o_{p^{\alpha+1}}(g) = \phi(p^{\alpha+1})$ and we are done. □

Lemma 2.56. *If p is an odd prime and $\alpha \in \mathbb{N}$, then there is a primitive root modulo $2p^\alpha$.*

Proof. Note that $\phi(2p^\alpha) = \phi(2)\phi(p^\alpha) = \phi(p^\alpha)$. By the above lemma, there is a primitive root g modulo p^α. Theorem 2.24 shows the existence of a number h such that

$$h \equiv 1 \mod 2, \quad h \equiv g \mod p^\alpha.$$

Clearly, h is coprime to $2p^\alpha$. Furthermore, since the congruence classes of the numbers

$$g^j, \quad 1 \le j \le \phi(p^\alpha)$$

are distinct modulo p^α, the congruence classes of the numbers

$$h^j, \quad 1 \le j \le \phi(p^\alpha) = \phi(2p^\alpha)$$

are distinct modulo $2p^\alpha$. This observation proves the lemma. □

Combining these lemmas gives Theorem 2.51. □

Next we discuss the problem of finding primitive roots when they exist. It is a consequence of Lemma 2.55 and Lemma 2.56 that once we know primitive roots modulo odd prime numbers, we can find primitive roots for odd prime powers and twice prime powers. The following lemma is easy to prove:

Lemma 2.57. *Let p be an odd prime number. Then a number g is a primitive root modulo p if and only if for all prime factors q of $p-1$, we have*

$$g^{\frac{p-1}{q}} \not\equiv 1 \mod p.$$

Proof. Suppose g is a primitive root. Then since for each prime factor q of $p-1$, $(p-1)/q < p-1$, we have $g^{(p-1)/q} \not\equiv 1 \bmod p$. For the other direction, if g is not a primitive root, then there is a divisor d of $p-1$ such that $1 < d < p-1$ and $g^d \equiv 1 \bmod p$. Let q be a prime divisor of $(p-1)/d$. Then $d \mid (p-1)/q$, which clearly implies $g^{(p-1)/q} \equiv 1 \bmod p$. □

As we will see momentarily this lemma gives a nice method to determine whether a given integer g is a primitive root modulo a prime number p, provided that $p-1$ has easily detectable prime factors. This can be a real challenge for a randomly chosen large prime number p. See the Notes at the end of this chapter for some comments on how this idea has been applied to cryptography.

Example 2.58. In this example we will use the lemmas proved above to determine primitive roots for the moduli $n = 17^\alpha, 2 \cdot 17^\alpha$. The proofs of Lemma 2.55 and Lemma 2.56 show that the key step is to find a primitive root modulo 17. In order to apply Lemma 2.57, we note $17 - 1 = 2^4$. Since the only prime factor of $17-1$ is 2, and $(17-1)/2 = 8$, Lemma 2.57 says that an integer g is a primitive root modulo 17 if and only if $17 \nmid g$ and $g^8 \not\equiv 1 \bmod 17$. The easiest way to search for candidates is by testing natural numbers in order starting with 2, jumping over squares. In our case, it is easy to check that

$$2^8 = 256 \equiv 1 \mod 17,$$

so 2 is not a primitive root modulo 17. Next, we check $g = 3$. We have

$$3^8 \equiv 16 \mod 17.$$

Hence $g = 3$ is indeed a primitive root modulo 17. Next, we check to see if $g = 3$ is a primitive root modulo 17^2. By the proof of Lemma 2.55, since $3^{16} \equiv 171 \not\equiv 1 \mod 17^2$, $g = 3$ is a primitive root modulo 17^2, and consequently a primitive modulo 17^α for all $\alpha \in \mathbb{N}$. Also, since $3 \equiv 1 \mod 2$, the proof of Lemma 2.56 implies that $g = 3$ is also a primitive root for $2 \cdot 17^\alpha$ for every $\alpha \in \mathbb{N}$.

Example 2.59. Using the method of the above example one can show that $g = 2$ is a primitive root modulo 19^α for every $\alpha \in \mathbb{N}$. Note that in this case since $19 - 1 = 2 \cdot 3^2$, a number g is a primitive root modulo 19 if and only if $g^9 \not\equiv 1 \mod 19$ and $g^6 \not\equiv 1 \mod 19$. Since 2 is even, it cannot be a primitive root modulo $2 \cdot 19^\alpha$ for any α. In this case $g = 2 + 19^\alpha$ is a primitive root modulo $2 \cdot 19^\alpha$ for all α.

Exercises

2.1 Prove Lemma 2.5.
2.2 Show that the alternative definitions in Definition 2.10 are equivalent.
2.3 Use the Euclidean Algorithm to give another proof for Theorem 2.12.
2.4 Prove Proposition 2.20.
2.5 Prove Proposition 2.21.
2.6 For the following pairs of integers (a, b), find integers x, y such that $\gcd(a, b) = ax + by$:

a. $(13, 15)$;
b. $(398, 270)$;
c. $(162, 65)$.

2.7 (✠) Find the gcd of 6437 and 12675. Find integers x, y such that $6437x + 12675y = \gcd(6437, 12675)$.
2.8 (✠) Find the gcd of 2594876242943772804330 and 11446995929696298.
2.9 Write the following number as a fraction $\frac{a}{b}$ with $a, b \in \mathbb{N}$ and $\gcd(a, b) = 1$:

$$10^{59} \left(\frac{1025}{1024}\right)^5 \left(\frac{1048576}{1048575}\right)^8 \left(\frac{6560}{6561}\right)^3 \left(\frac{15624}{15626}\right)^8 \left(\frac{9801}{9800}\right)^4.$$

Determine the prime factorizations of a, b without the use of a computer. Mossaheb [34] attributes this problem to Gauss.
2.10 Determine all natural numbers n such that $\prod_{d|n} d = n^2$.
2.11 Suppose for integers a, m, n, k we have $a^m \equiv 1 \mod k$ and $a^n \equiv 1 \mod k$. Show that $a^{\gcd(m,n)} \equiv a^{\operatorname{lcm}(m,n)} \equiv 1 \mod k$.
2.12 Show that if a rational number $\frac{a}{b}$, with $a, b \in \mathbb{Z}$ and $\gcd(a, b) = 1$, satisfies the equation

$$a_n x^n + a_{n-1} x^{n-1} + \dot{+} a_1 x + a_0 = 0,$$

with $a_0, a_1, \ldots, a_n \in \mathbb{Z}$, then $a \mid a_0$ and $b \mid a_n$. Use this result to find the rational roots of the following equations:

a. $5x^3 + 8x^2 + 6x - 4 = 0$;
b. $x^5 - 7x^3 - 12x^2 + 6x + 36 =$;
c. $6x^6 - x^5 - 23x^4 - x^3 - 2x^2 + 20x - 8 = 0$.

2.13 Use the previous exercise to show $\sqrt{2} + \sqrt[3]{3}$ is irrational.
2.14 Show that for all integers a, b, c, d satisfying $ad - bc = 1$ we have $\gcd(a + b, c + d) = 1$.
2.15 Show that for all integers $n > 1$, $1 + 1/2 + 1/3 + \cdots + 1/n$ is not an integer.
2.16 If f is a non-constant polynomial with integer coefficients, $f(n)$ is composite for infinitely many values of n.
2.17 Show that if p, q are prime numbers larger than 3, then the remainder of division of $p^2 + q^2$ by each of the numbers 3, 4, 6, 12, and 24 is equal to 2.
2.18 (✠) Show that a number n is prime if and only if it is not divisible by any natural numbers m with $1 \leq m \leq n^{1/2}$. This result is known as the *Sieve of Eratosthenes*. Use this idea to list all prime numbers between 1 and 1000.
2.19 (✠) Find five natural numbers k such that $22 + 37k$ is a prime number.
2.20 Show that for all $m, n \in \mathbb{N}$,

$$\frac{\gcd(m, n)}{n} \binom{n}{m}$$

is an integer.
2.21 Suppose $F_n = 2^{2^n} + 1$. Show that for all $m > n$, $F_n \mid F_m - 2$.
2.22 Find necessary and sufficient conditions for the solvability of the system (2.4). Find the general solution of the system.
2.23 Solve the system of congruence equations

$$\begin{cases} 3x \equiv 1 \mod 4, \\ 3x \equiv 1 \mod 13, \\ 5x \equiv 11 \mod 21. \end{cases}$$

2.24 Find the general integral solution of the Diophantine equation

$$239x - 111y = 1.$$

2.25 Find all pairs of integers (x, y) satisfying the equation $6x + 9y = 12$.
2.26 (✠) Find all x such that $85x \equiv 970 \mod 64322$.
2.27 (✠) Find all solutions of $37x \equiv 217 \mod 8600$.
2.28 (✠) Find all x that satisfy the following system of congruence equations:

$$\begin{cases} x \equiv 12 & \mod 64; \\ x \equiv 1 & \mod 173; \\ x \equiv 5 & \mod 715. \end{cases}$$

2.29 Show that $5!25! \equiv 1 \mod 31$.

2.30 Show that if $p \equiv 3 \mod 4$, then

$$\left((\frac{p-1}{2})! \right)^2 \equiv 1 \mod p.$$

2.31 Prove the uniqueness assertion of Lemma 2.37.

2.32 Give two different proofs for the statement that for all integers n, $n^5/5+n^3/3+7n/15 \in \mathbb{Z}$. Generalize.

2.33 Find all the solutions to the congruence $x^2 \equiv 1 \mod 264$.

2.34 By examining the solutions of the equation $x^2 \equiv 1 \mod p$, show that for all primes p, $(p-1)! \equiv -1 \mod p$. Show that if $n > 4$ is not prime, $(n-1)! \equiv 0 \mod n$.

2.35 Let $n \in \mathbb{N}$. Compute the product

$$\prod_{\substack{1 \le d \le n \\ d^2 \equiv 1 \mod n}} d.$$

Use your formula to determine

$$\prod_{\substack{1 \le d \le n \\ \gcd(d,n)=1}} d.$$

2.36 Find the roots of the polynomials $x^2 - x + 1$ and $x^2 - x + 2$ modulo 7.

2.37 (✠) Find an integer x such that $x^2 \equiv 1879121 \mod 3698963$.

2.38 (✠) Find the last four digits of 2^{4000}.

2.39 Show that $\phi(n)$ is even if and only if $n > 2$.

2.40 Determine all n such that $\phi(n) = 6$.

2.41 Determine all n such that $\phi(n) = 40n/77$.

2.42 Show that for every odd integer $n > 1$ we have $\phi(n) > \sqrt{n}$.

2.43 Determine all n with $\phi(n) \mid n$.

2.44 Give a different proof for Theorem 2.32 using the *Inclusion–Exclusion Principle*.

2.45 Prove the following generalization of Theorem 2.32: If $\gcd(m, n) = d$, then

$$\phi(mn) = \phi(m)\phi(n)\frac{d}{\phi(d)}.$$

2.46 Use Theorem 2.33 to give another proof for Theorem 2.34.

2.47 Show that for each complex number $\alpha \neq 1$ and for each natural number n, we have

$$\sum_{k=0}^{n-1} \alpha^k = \frac{1-\alpha^n}{1-\alpha};$$

Show that if $|\alpha| < 1$, then

$$\sum_{k=0}^{\infty} \alpha^k = \frac{1}{1-\alpha}.$$

2.48 Show that for each $g > 4$, the number $(4.41)_g$ is the square of a rational number. Find its square root. Repeat the same problem for $(148.84)_g$ for $g > 8$.
2.49 For what values of g, are the numbers $(0.16)_g$, $(0.20)_g$, $(0.28)_g$ the consecutive terms of a geometric sequence?
2.50 If $25/128 = (0.0302)_g$, find g.
2.51 Find the base 5 expansion of $2877/3125$.
2.52 Find the base 9 expansion of $(200.211)_3$.
2.53 Determine the rational number with base 7 expansion $(0.\overline{130})_7$. Solve the same problem for $(0.1\overline{296})_{12}$.
2.54 Find all primitive roots modulo 38.
2.55 (✠) Find all primes $p < 1000$ for which 3 is a primitive root.
2.56 (✠) Find five primitive roots modulo 100003.
2.57 (✠) Find five primitive roots modulo 987654103^2.
2.58 If $p = 4q + 1$ and $q = 3r + 1$ are prime, show that 3 is a primitive root modulo p.
2.59 Let p, q be distinct primes. Find the number of solutions of $x^p \equiv 1 \bmod q$ in terms of p, q.
2.60 Without using a computer, prove that $2^{17} - 1$ is prime. Hint: Show that if it is not prime, it must be divisible by one of the numbers 103, 137, 239, or 307.
2.61 Let p be an odd prime such that $(p-1)/2$ is an odd prime. Prove that if a is a positive integer with $1 < a < p-1$, then $p - a^2$ is a primitive root modulo p.
2.62 Show that n is a square if and only if $d(n)$ is odd.
2.63 Show that for all $n, s, t \in \mathbb{N}$, with $s \neq t$, $s - t \mid d(n^s) - d(n^t)$.
2.64 Show that for all $m, n, s \in \mathbb{N}$, we have $s \mid d(m^s) - d(n^s)$.
2.65 Find necessary and sufficient conditions on integers a, b, c, d so that there are integers x, y, z satisfying the system of Diophantine equations

$$\begin{cases} 4x + y + az = b, \\ x + y + cz = d. \end{cases}$$

2.66 Let p be an odd prime. Show that if we write $1 + 1/2 + \cdots + 1/(p-1)$ as a fraction a/b with $a, b \in \mathbb{N}$, then $p \mid a$. A far more interesting problem is to show that if $p \geq 5$, $p^2 \mid a$.

Notes

Historical references

The standard reference for the history of classical number theory is Dickson's *History of the theory of numbers* in three volumes. Most of the material in this chapter has been reviewed in the first volume [15], especially Ch. III, V, and VIII. A more current reference for the history of mathematics is [9]. As impressive as these books are, like many other books on the history of science, they are unfortunately very Eurocentric. The history of mathematics as told through these and other similar texts runs like this: The Greeks invented mathematics; then as Europe was falling into the Dark Ages, Muslims ran to the rescue; Muslims carefully guarded mathematics for a few centuries; with the arrival of the Renaissance, the Muslims handed mathematics back to the Europeans who gracefully accepted the gift, and who have ever since been championing the progress of mathematics. This Eurocentricity does not stop at the history, and in fact it permeates every aspect of the practice of mathematics. In reality the history of mathematics is far more complicated and far more multicultural than a simple straight line connecting Athens of the antiquity to the North America and Europe of 21st century.

In this book I have made a conscious effort to highlight contributions by non-Europeans to number theory. However—and this is far from an acceptable excuse—because of my lack of expertise as well my own Eurocentric education I am not able to do justice to the subject. Getting the history right is not just a matter of intellectual curiosity. Those of us who work as educators in North America are acutely aware of the fact that a good portion of our students are not of European descent. To many of our students mathematics is a European invention, and will continue to be practiced by Europeans and people of European descent. Nothing could be further from the truth. Mathematics has been practiced on every continent, by all sorts of people, for thousands of years, and there are distinguished mathematicians of every imaginable background today doing fantastic mathematics—and this should be emphasized in our teaching. There is, unfortunately, a shortage of modern, easily accessible texts putting in the correct historical perspective the progress of mathematics through the millennia. Even in cases where a serious mathematician such as van der Waerden [55] has attempted to write a history of mathematics as inspired by the progress made by non-Europeans, the works of these non-Europeans are described in relation to and within the framework of modern European mathematics, or the Greek mathematics of the antiquity, in the sense that, what of the works of non-Europeans that has not been superseded and swallowed by some mathematical work developed by a European mathematician is often not considered worthy of review. The same problem exists in most works written by European or North American historians of mathematics, with a notable exception being Plofker [39]. Writings by historians like Roshdi Rashed, especially the second volume of *Encyclopedia of the history of Arabic sciences* [40] which covers mathematics, and Joseph [28] are good alternatives to standard Eurocentric narratives that saturate the literature.

On a personal note, growing up in Iran, I never felt that mathematics was a European invention or practice—I knew of Iranian mathematicians like Omar Khayyam, Mohammad Al-Khwarizmi, and Mohammad Karaji, and these were people I identified with. I credit Iranian education pioneers like G. H. Mossahab, M. Hashtroodi, M. Hessabi in the 1940s and 1950s, and more recently S. Shahshahani, P. Shahriari, O. A. Karamzadeh, Y. Tabesh, and others starting in the 1970s, for initiating the effort to instill the notion in the minds of the Iranian youth that mathematics, along with other sciences, was as Iranian as apple pie is American. It is because of their efforts that Iran has enjoyed a revitalization of mathematics in the last 25 years. Culture building takes time, and, as in the case of those Iranian pioneers, one may not live long enough to see the fruits of one's labor, but with patience and perseverance great things are possible.

Euclid and his Elements

Euclid (325–265 BCE) was the person who transformed mathematics from a number of uncoordinated and loosely proven theorems into an articulated and surely grounded science. Some of the theorems in Euclid's Elements were previously known by other mathematicians: Thales (624–546 BCE) who was according to Aristotle the first Greek philosopher, Eudoxus (410–355 BCE), Pythagoras and other Pythagoreans, etc. A predecessor to Euclid was Hippocrates (470-410 BCE) who wrote the first *Elements* around 430 BCE. Euclid was extremely rigorous in his treatment of mathematics. (Though as noted by David Hilbert [26], Euclid should have augmented his postulates by adding a few more.) E. T. Bell argues that if the world had followed Archimedes as opposed to Euclid, Calculus would have been discovered before the birth of Christ. This is a harsh criticism of the Euclidean rigor, and of the course of history, but it is nonetheless most likely true that the sort of rigor that Euclid brought into mathematics slowed down progress in some sense. Archimedes was a master problem solver who was interested in the applications of mathematics in the real world. Euclid, on the other hand, was interested in gaining a deep understanding of concepts via systematic study. For what it is worth, almost 2500 years later, we still practice mathematics the way Euclid did mathematics in his magnum opus. An interesting feature of *the Elements* is that the writing is extremely homogeneous. Euclid makes no distinction between trivial facts and deep theorems, and everything is proved with the same degree of care. Was Euclid really not aware that some of his results are more important than others? We will never know.

The theory of numbers is treated by Euclid in books 7-10 of the *Elements*. At the beginning of Book 7 Euclid lists definitions: unit, numbers, multiple, even and odd number, prime and composite numbers, square, proportional, perfect number, etc. These are very much in the Pythagorean style, but with some modifications. We refer the reader to the excellent commentary in Sir Thomas L. Heath's "The Thirteen Books of Euclid's Elements" [20] published in 1926. In this book Sir Heath compares Euclid's definitions to those given by his predecessors. In the case of prime numbers, Euclid's definition varies slightly from the one written by the Pythagorean Philolaus

(480–390 BCE) who seems to have been the first person to give a definition of prime numbers.

For all their aura of naturalness, prime numbers almost never appear in nature for reasons of primality. The only example of such a process is the life cycles of a certain genus of cicadas. These insects spend most of their lives underground and emerge to daylight every 13 or 17 years. The fact that 13 and 17 are prime numbers gives these insects a computable but small evolutionary edge over their predators. Over millions of years the evolutionary edge of these insects has helped them not go extinct. Beyond this, we are not aware of any cosmic or earthly processes that produce prime numbers for reasons of primality. Even within mathematics, as practiced by human beings, it appears that prime numbers were an invention of the Greeks, and that no one else in the ancient world had a notion of prime numbers. Mathematicians in Babylon, India, China, and the Americas investigated very sophisticated mathematical theories, including those applicable to astronomy and other sciences, but as far as we can tell none of these mathematicians had a theory of prime numbers.

For more on Euclid's work on prime numbers, see Notes, Chapter 6.

Natural Numbers and mathematical induction

In this book we will treat natural numbers in a common sense, intuitive fashion. We assume the set of natural numbers $\mathbb{N}$ consists of positive integers $1, 2, 3, \ldots$, equipped with the standard addition and multiplication operations, enjoying the familiar properties of commutativity and associativity for addition and multiplication, and distribution laws for multiplication over addition. We also *know* that we can prove statements in the set of natural numbers using *mathematical induction*, accepted as an axiom. In reality, however, all of these statements are non-trivial and require close examination. The axiomatic study of the set of natural numbers has a long, rich history. We refer the reader to [18, Ch. 1] for an accessible introduction to this beautiful subject.

Number-theory-based cryptography

Many modern cryptographic methods are based on the material presented in this chapter. Here we will explain two standard techniques. For an elementary treatment of these methods and other number theoretic cryptosystems we refer the reader to [53].

The RSA Cryptosystem, named after Ron Rivest, Adi Shamir, and Leonard Adleman, is based on the notion that while multiplying numbers is easy, finding the prime factors of a large number is difficult. More specifically, if we know the prime factors of a natural number n, then Theorem 2.33 tells us how to compute the value of $\phi(n)$. However, without knowing the prime factors of n, we do not have a fast algorithm to

compute $\phi(n)$. Presumably, one can take (2.6) as the definition of $\phi(n)$. This requires going through the list of numbers 1 to n and examining the gcd of each one with n, which, if the number n is of the order of 10^{500}, would be impossible.

RSA is an example of a *public key cryptosystem*. In such a cryptographic scheme an individual A sets up a public key K, which is available to everyone, and keeps a private piece of information S, which is kept secret. The idea is that anyone who wants to communicate with A will encrypt the message using the publicly available key K but decrypting the encrypted message requires the secret information S. In the case of RSA, the public key is a large natural number n which is the product of prime numbers p, q. The prime numbers p and q are kept secret..

This is how RSA works. Suppose Azadeh wants to set up a public key. She picks large prime numbers p, q. She computes $n = pq$, $\phi(n) = (p-1)(q-1)$, and she picks a natural number e such that that $\gcd(e, \phi(n)) = 1$. She also finds an integer d such that $ed \equiv 1 \bmod \phi(n)$, i.e, $ed = 1 + u\phi(n)$ for some integer u. She will keep p, q, d, and $\phi(n)$ secret, but publishes the pair (n, e). Now suppose Azadeh's friend, Behnam, wants to communicate with Azadeh. Suppose the message that Behnam wants to send has numerical value m, obtained using ASCII or some other method (technically speaking, Behnam will have to make sure that $\gcd(m, n) = 1$). Behnam downloads the pair (n, e) from Azadeh's public profile, and computes $y := m^e \bmod n$, i.e., the remainder of the division of m^e by n which will be a number between 0 and n. Behnam keeps the message m secret, but sends the message y to Azadeh over some public channel, e.g., Facebook or SMS. Azadeh receives the message y, and deciphers it by computing

$$y^d \equiv (m^e)^d \equiv m^{1+u\phi(n)} \equiv (m^{\phi(n)})^u \cdot m \equiv m \mod n,$$

after using Theorem 2.31. On the other hand, Esmat, an evil person, is listening to the conversation happening between Azadeh and Behnam. Esmat downloads the message y. She also knows (n, e) as these are publicly available. However, at present there is no reasonably fast way to get from the data y, (n, e) to m without knowing d, and knowing d requires $\phi(n)$. As noted above computing $\phi(n)$, at the time of this writing, requires knowing the prime factors of n, which Azadeh is keeping secret.

For example, suppose Azadeh picks the prime numbers $p = 101$ and $q = 113$ (this is just a prototype; in practice the prime numbers are a few hundred digits long). Hence $n = 101 \times 113 = 11413$. We have $\phi(n) = (101-1)(113-1) = 11200$. She also picks $e = 3$. Note that $\gcd(3, 11200) = 1$. Azadeh's public key is the pair $(11413, 3)$. What Azadeh is not sharing with the public are the prime numbers 101 and 113. She also keeps secret the number d such that $3d \equiv 1 \mod 11200$. Azadeh can easily compute, for example using SageMath, Appendix C, that $d = 7467$ works. Now suppose Behnam wants to transmit a message m with numerical value 77 to Azadeh. Behnam computes $m^e \bmod n$. In this case since $m = 77$, $e = 3$, and $n = 11413$, he computes

$$77^3 \equiv 13 \mod 11413.$$

So Behnam's message, which he can communicate over a public channel, is $y = 13$. Anyone can read this message x, and everyone knows Azadeh's public key

(11413, 3). So the problem that Esmat, the evil person, needs to solve is this: Find m such that $m^3 \equiv 13 \mod 11413$. For Azadeh, this is easy. All she needs to do is compute

$$13^{7467} \equiv 77 \mod 11413,$$

which she can easily do using SageMath.

The ElGamal Cryptosystem, named after the Egyptian computer scientist Taher ElGamal, is based on the difficulty of the Discrete Log problem. As mentioned earlier RSA cryptography is based on the idea that it is difficult to go from $(m^e \bmod n, e, n)$ to m. The flip side of this idea is the *Discrete Log* problem. Let n be a natural number for which we have a primitive root g. Let $1 < x < n$ be a natural number that is coprime to n. The *Discrete Log* problem asks for the determination of an integer $0 < l < \phi(n)$ such that $x \equiv g^l \bmod n$.

In the ElGamal Cryptosystem, Azadeh picks a large prime p, a primitive root g modulo p, a random number l, with $1 < l < p-1$, and computes $e = g^l \bmod n$. Azadeh's public key is (p, g, e) which she publishes. She keeps l secret. Behnam wants to send a message m to Azadeh. Benham picks a random integer u, $1 < u < p-1$, and computes $x := g^u \bmod p$, and $y := m \cdot e^u \bmod p$. Behnam sends the pair (x, y) over a public channel to Azadeh. Azadeh recovers m by computing

$$y \cdot x^{-l} \equiv m \cdot (g^l)^u \cdot (g^u)^{-l} \equiv m \mod p.$$

We refer the reader to [53], especially Ch. 6 for RSA and Ch. 7 for ElGamal.

Primitive roots and Artin's conjecture

The notion of the order of a modulo n made an appearance in Gauss's book [21, articles 315-317], when he considered the decimal expansion of $1/p$ for a prime number p, $p \neq 2, 5$. In this case, the fraction $1/p$ is purely periodic and its period is equal to $o_p(10)$. In general, we saw in this chapter that if m, n are natural numbers with $\gcd(m, n) = 1$, then the base n expansion of $1/m$ is purely periodic with minimal period equal to $o_m(n)$. In particular the minimal period is at most equal to $\phi(m)$. In the case where $m = p$ is a prime number, $\phi(p) = p-1$. The following is a natural question: For a natural number n, are there infinitely many prime numbers p such that the base n-expansion of $1/p$ has period $p-1$? Note that in this case n will have to be a primitive root modulo p. While the answer to the question is expected to be yes, it is not known for any n, not even $n = 10$.

Conjecture 2.60 (Artin 1927). Fix an integer $g \neq -1, 0, 1$ which is not a perfect square. Then there are infinitely many primes p such that g is a primitive root modulo p.

In fact, Artin conjectured an asymptotic formula for $\#\{p \text{ prime} \mid p \leq X, o_p(g) = p-1\}$ of the form $\delta(g)X/\log X$ as $X \to \infty$, for some constant $\delta(g) > 0$. Artin gave a heuristic argument to derive a formula for $\delta(g)$; however, in 1957 Derrick and Emma

Lehmer observed that Artin's predicted formula did not match numerical data. Artin was then able to pinpoint the error in the original heuristic reasoning and corrected the prediction. In 1967 Hooley [81] gave a proof of the predicted asymptotic formula which relied on some version of Riemann's Hypothesis, not yet proved; see Notes to Chapter 13. See Murty's expository article [89] for an accessible account of the progress made toward the conjecture up until the time of its publication. For a more up-to-date report on the conjecture and the methods and techniques used in its study, see Moree's survey [88].